MATHEMATISCHE ZEITSCHRIFT.

UNIVERSITY OF HAWAII LIBRARY

UNTER STÄNDIGER MITWIRKUNG

VON

K. KNOPP
TÜBINGEN

E. SCHMIDT
BERLIN

I. SCHUR
BERLIN

HERAUSGEGEBEN

VON

L. LICHTENSTEIN
LEIPZIG

WISSENSCHAFTLICHER BEIRAT:

L. BIEBERBACH W. BLASCHKE L. FEJÉR G. H. HARDY E. HECKE
G. HERGLOTZ E. LANDAU O. PERRON F. SCHUR H. WEYL

34. BAND

QA1
M88
v. 34

BERLIN
VERLAG VON JULIUS SPRINGER
1932

5856-42

MATHEMATISCHE ZEITSCHRIFT

UNTER STÄNDIGER MITWIRKUNG

VON

K. KNOPP
TÜBINGEN

E. SCHMIDT
BERLIN

I. SCHUR
BERLIN

HERAUSGEGEBEN

VON

L. LICHTENSTEIN
LEIPZIG

WISSENSCHAFTLICHER BEIRAT:

L. BIEBERBACH W. BLASCHKE L. FEJÉR G. H. HARDY E. HECKE
G. HERGLOTZ E. LANDAU O. PERRON F. SCHUR H. WEYL

34. BAND

BERLIN
VERLAG VON JULIUS SPRINGER
1932

66 52915

QA1
M88
V.34

Inhaltsverzeichnis

Seite

Ayres, W. L., On avoidable Points of Continua with an Application to end Points 161

Baer, R., und Fr. Levi, Stetige Funktionen in topologischen Räumen. Ernst Zermelo zum 60. Geburtstag am 27. Juli 1931 gewidmet 110

Barth, W., Bemerkungen zu der Zuschrift der Herren J. Bonder und S. Neumark in Band 32 dieser Zeitschrift 279

Boerner, H., Das Eigenwertproblem der selbstadjungierten linearen Differentialgleichung vierter Ordnung 293

Bochner, S., Ein Konvergenzsatz für mehrvariablige Fouriersche Integrale . . . 440

Breusch, R., Zur Verallgemeinerung des Bertrandschen Postulates, daß zwischen x und $2x$ stets Primzahlen liegen 505

Burkill, J. C., The Approximately Continuous Perron Integral 270

Carathéodory, C., Über die Abbildungen, die durch Systeme von analytischen Funktionen von mehreren Veränderlichen erzeugt werden 758

Chwistek, L., Neue Grundlagen der Logik und Mathematik. Zweite Mitteilung 527

Egerváry, E., Verschärfung eines Harnackschen Satzes und anderer Abschätzungen für nichtnegative harmonische Polynome 741

Evelyn, C. J. A., and E. H. Linfoot, On a Problem in the additive Theory of Numbers. (Third paper) 637

Hardy, G. H., and J. E. Littlewood, An additional note on Parseval's Theorem 634

Hardy, G. H., and J. E. Littlewood, Some properties of fractional integrals. II. 403

Hardy, G. H., and J. E. Littlewood, Some new cases of Parseval's Theorem . 620

Heffter, L., Über den Cauchyschen Integralsatz. (Nachtrag zu der Arbeit Band 32, S. 476—480) 473

Hlavatý, V., Projektive Invarianten einer Kurvenkongruenz und einer Kurve . 58

Hopf, E., Über den funktionalen, insbesondere den analytischen Charakter der Lösungen elliptischer Differentialgleichungen zweiter Ordnung 194

v. Ignatowsky, W., Über doppelpolige Lösungen der Wellengleichung 1

Kapferer, H., Über die Jacobische Kurve eines Systems von drei Kurven und den Begriff „Abhängigkeit" bei homogenen Polynomen 554

Karamata, J., Über einen Satz von Vijayaraghavan 737

Kerner, M., Minimaxeigenschaft der offenen und geschlossenen Extremalen . . 645

Lense, J., Über die Ableitungsgleichungen der ametrischen Mannigfaltigkeiten . 721

Levi, Fr., und R. Baer, Stetige Funktionen in topologischen Räumen. Ernst Zermelo zum 60. Geburtstag am 27. Juli 1931 gewidmet 110

Lösch, F., Über den Permanenzsatz gewisser Limitierungsverfahren für Doppelfolgen 281

Lösch, F., Über das Verhalten der Potenzreihen auf dem Rande des Konvergenzkreises II. 291

Linfoot, E. H., and C. J. A. Evelyn, On a Problem in the additive Theory of Numbers. (Third paper) 637

Seite

Littlewood, J. E., and G. H. Hardy, Some properties of fractional integrals. II. 403
Littlewood, J. E., and G. H. Hardy, Some new cases of Parseval's Theorem . 620
Littlewood, J. E., and G. H. Hardy, An additional note on Parseval's Theorem 634
Maier, W., Orthogonale Systeme elliptischer Funktionen 487
Mitteilungen 320
Nagell, T., Sätze über algebraische Ringe 179
Nagell, T., Zur algebraischen Zahlentheorie 183
Nikliborc, W., Über die Abplattung der homogenen Gleichgewichtsfiguren rotierender, gravitierender Flüssigkeiten 74
Püschel, W., Die erste Randwertaufgabe der allgemeinen selbstadjungierten elliptischen Differentialgleichung zweiter Ordnung im Raum für beliebige Gebiete 535
Reichenbach, H., Axiomatik der Wahrscheinlichkeitsrechnung 568
Ridder, J., Über den Perronschen Integralbegriff und seine Beziehung zu den R-, L- und D-Integralen 234
Romanovski, P., Quelques considérations sur la théorie des intégrales singulières 35
Salié, H., Über die Kloostermanschen Summen $S(u, v; q)$ 91
Schilling, F., Die numerische Lösung der Gleichungen vierten Grades 50
Sidon, S., Einige Sätze und Fragestellungen über Fourier-Koeffizienten 477
Sidon, S., Ein Satz über Fouriersche Reihen mit Lücken 481
Sidon, S., Ein Satz über die Fourierschen Reihen stetiger Funktionen 485
Süß, W., Über wechselseitig affinparallele Flächen 158
Thomsen, G., Über einen neuen Zweig geometrischer Axiomatik und eine neue Art von analytischer Geometrie 668
Walfisz, A., Teilerprobleme. Zweite Abhandlung 448
Weber, W., Umkehrbare Ideale 131
Wintner, A., Grundlagen einer Genealogie der periodischen Bahnen im restringierten Dreikörperproblem. Erste Mitteilung. Beweis des E. Strömgrenschen dynamischen Abschlußprinzips der periodischen Bahngruppen 321
Wintner, A., Grundlagen einer Genealogie der periodischen Bahnen im restringierten Dreikörperproblem. Zweite Mitteilung. Sortengenealogie, Hekubakomplex und Gruppenfortsetzung 350

Über doppelpolige Lösungen der Wellengleichung.

Von

W. v. Ignatowsky in Leningrad.

Inhaltsverzeichnis.

		Seite
Einleitung		1
§ 1.	Elliptische Koordinaten	2
§ 2.	Einführung neuer Koordinaten	3
§ 3.	Beziehungen für die neuen Koordinaten und Bemerkungen	4
§ 4.	Verschiedene Typen von Integralen der Wellengleichung	8
§ 5.	Funktionentheoretische Betrachtungen	11
§ 6.	Doppelpolige Funktionen	16
§ 7.	Analytische Untersuchungen der Funktionen des § 5	17
§ 8.	Einige Folgerungen und Einführung von doppelpoligen Integralen der Wellengleichung	21
§ 9.	Weitere Folgerungen und Zusammenstellung	24
§ 10.	Übergang zu $a = 0$	26
§ 11.	Beziehungen zwischen verschiedenen Integralen der Wellengleichung und Folgerungen	29
§ 12.	Betrachtungen zur Beugung an einem blanken und schwarzen Streifen	30

Einleitung.

Die exakte Lösung des Problems der Beugung an einer Scheibe und einem Streifen begegnet mannigfachen formellen Schwierigkeiten, auf die ich verschiedentlich schon hingewiesen habe[1]); am ehesten versprach die

[1]) W. v. Ignatowsky, Zur Theorie der Beugung an schwarzen Schirmen und Erwiderung an F. Kottler, Ann. d. Physik 77 (1925), Anm. zu § 9, und: Zur Beugungstheorie, Journ. d. Russ. phys. u. chem. Ges., phys. Teil, 56 (1924), S. 613 (russisch); (Physik. Berichte 1925, S. 473).

Methode einen Erfolg, auf die a. a. O. Sommerfeld[2]) hingewiesen hat und die auf der Bildung von „doppelpoligen“ Lösungen der Wellengleichung — mit zwei im Endlichen gelegenen Verzweigungspunkten — beruhte.

Die vorliegende Arbeit führt nun diesen Gedanken aus. Es werden besondere Koordinaten eingeführt (§ 2), worauf die umgeformte Wellengleichung in der angedeuteten Weise gelöst werden kann (§ 5). Die Einführung zweiwertiger oberen Grenzen bei den auftretenden Integralen, in Verbindung mit einer weiteren Zweiwertigkeit unter dem Integralzeichen selbst, erfordert eingehendere funktionentheoretische Betrachtungen (§ 5), die erst die analytischen Entwicklungen (§ 7) ermöglichen. Es können daraufhin (§ 8) explizite doppelpolige Integrale der Wellengleichung ausgeschrieben werden, die für die Lösung des Beugungsproblems wenigstens das formale Rüstzeug liefern.

§ 1.

Elliptische Koordinaten.

Die Wellengleichung lautet im ebenen Fall:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \varkappa^2 u = 0, \tag{1}$$

wobei $\varkappa = \frac{2\pi}{\lambda}$ und λ die Wellenlänge bedeuteten.

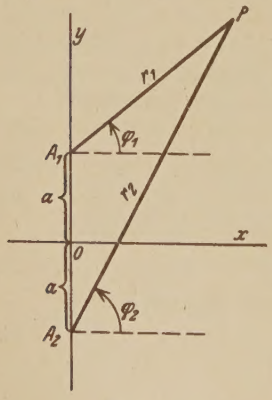


Fig. 1.

Die Lagen der X- und Y-Achsen sind aus der Fig. 1 ersichtlich und die Z-Achse ist zum Beschauer hin gerichtet.

Wir führen nun vorläufig die elliptischen Koordinaten ξ und δ ein, welche wir folgendermaßen definieren:

$$x = a \sinh \xi \cdot \cos \delta \tag{2}$$

und

$$y = a \cosh \xi \cdot \sin \delta. \tag{3}$$

Auf Grund dieser Koordinaten erhalten wir statt (1):

$$\frac{\partial^2 u}{\partial \xi^2} + \frac{\partial^2 u}{\partial \delta^2} + q^2 u \{\cosh^2 \xi - \sin^2 \delta\} = 0, \tag{4}$$

wobei $q = \varkappa a$ bedeuten soll.

Ferner haben wir

$$\frac{\partial}{\partial x} = \frac{1}{a(\cos^2 \delta + \sinh^2 \xi)} \left\{ \cos \delta \cdot \cosh \xi \cdot \frac{\partial}{\partial \xi} - \sin \delta \cdot \sinh \xi \cdot \frac{\partial}{\partial \delta} \right\} \tag{5}$$

und

$$\frac{\partial}{\partial y} = \frac{1}{a(\cos^2 \delta + \sinh^2 \xi)} \left\{ \sin \delta \cdot \sinh \xi \cdot \frac{\partial}{\partial \xi} + \cos \delta \cdot \cosh \xi \cdot \frac{\partial}{\partial \delta} \right\}. \tag{6}$$

[2]) A. Sommerfeld, Mathematische Theorie der Diffraction, Math. Annalen 47 (1896), S. 374.

Außerdem wird sein (Fig. 1)

$$r_1 = a\{\cosh\xi - \sin\delta\} \tag{7}$$

und

$$r_2 = a\{\cosh\xi + \sin\delta\}. \tag{8}$$

§ 2.

Einführung neuer Koordinaten.

Wir führen nun folgende Paare von Koordinaten ein

$$w_1 = \cosh\left\{\frac{\xi + i\left(\delta + \frac{\pi}{2}\right)}{2}\right\}; \qquad t_1 = \cosh\left\{\frac{\xi - i\left(\delta + \frac{\pi}{2}\right)}{2}\right\} \tag{1}$$

und

$$w_2 = \sinh\left\{\frac{\xi + i\left(\delta + \frac{\pi}{2}\right)}{2}\right\}; \qquad t_2 = \sinh\left\{\frac{\xi - i\left(\delta + \frac{\pi}{2}\right)}{2}\right\}. \tag{2}$$

Hieraus folgt

$$w_1^2 = w_2^2 + 1; \quad t_1^2 = t_2^2 + 1, \tag{3}$$

und demnach

$$\begin{aligned} w_1 &= \sqrt{w_2^2 + 1}; \quad & t_1 &= \sqrt{t_2^2 + 1}, \\ w_2 &= \sqrt{w_1^2 - 1}; \quad & t_2 &= \sqrt{t_1^2 - 1}. \end{aligned} \tag{4}$$

Bezüglich der hier auftretenden Quadratwurzeln werden wir ausführlich im folgenden Paragraphen sprechen, vorläufig folgt aus (3) und (4)

$$\begin{aligned} \frac{\partial}{\partial w_1} &= \frac{\partial}{\partial w_2}\cdot\frac{w_1}{w_2}; & \frac{\partial}{\partial t_1} &= \frac{\partial}{\partial t_2}\cdot\frac{t_1}{t_2}, \\ \frac{\partial}{\partial w_2} &= \frac{\partial}{\partial w_1}\cdot\frac{w_2}{w_1}; & \frac{\partial}{\partial t_2} &= \frac{\partial}{\partial t_1}\cdot\frac{t_2}{t_1} \end{aligned} \tag{5}$$

und

$$\frac{\partial w_1^2}{\partial w_1} = \frac{\partial w_2^2}{\partial w_1} = 2w_1; \qquad \frac{\partial t_1^2}{\partial t_1} = \frac{\partial t_2^2}{\partial t_1} = 2t_1. \tag{6}$$

Auf Grund der Koordinaten (1) geht (4) § 1 über in

$$\frac{\partial^2 u}{\partial w_1\,\partial t_1} + 4q^2 w_1 t_1 u = 0. \tag{7}$$

Statt (7) können wir wegen (5) und (6) auch schreiben

$$\frac{\partial^2 u}{\partial w_2\,\partial t_2} + 4q^2 w_2 t_2 u = 0, \tag{8}$$

oder auch

$$\frac{\partial^2 u}{\partial w_2\,\partial t_1} + 4q^2 w_2 t_1 u = 0; \qquad \frac{\partial^2 u}{\partial w_1\,\partial t_2} + 4q^2 w_1 t_2 u = 0. \tag{9}$$

Da, wie wir später sehen werden, sich (1) nur auf den Pol A_1 und (2) sich nur auf den Pol A_2 bezieht, so folgt aus (7) bis (9), daß wir entweder jeden Pol einzeln oder beide zusammen behandeln können.

Um die Bedeutung der neuen Koordinaten klarer hervortreten zu lassen, gehen wir zu den parabolischen Koordinaten ζ und η über mit dem Brennpunkt z. B. im Pole A_1.

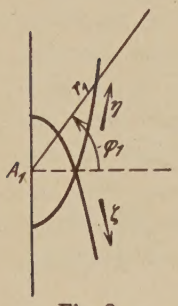


Fig. 2.

Dann haben wir bekanntlich (Fig. 2)

$$(10)\quad \zeta = \sqrt{2r_1}\cos\tfrac{1}{2}\left(\varphi_1+\tfrac{\pi}{2}\right);\quad \eta = \sqrt{2r_1}\sin\tfrac{1}{2}\left(\varphi_1+\tfrac{\pi}{2}\right).$$

Zwischen diesen Koordinaten und (1) bestehen die Beziehungen

$$(11)\qquad \zeta = \sqrt{a}(w_1+t_1);\qquad \eta = -i\sqrt{a}(w_1-t_1),$$

aus welchen klar hervorgeht, daß (1) tatsächlich sich nur auf den Pol A_1 bezieht.

Wir wollen sofort auf eine wichtige und grundlegende Eigenschaft der Gleichungen (7) bis (9) aufmerksam machen.

Es seien $u(w_1, t_1)$, $u_1(w_2, t_1)$ usw. Integrale dieser Gleichungen.

Dann sind auch, wie man sich leicht überzeugen kann, $u(\sqrt{w_1^2+c^2}, \sqrt{t_1^2+e^2})$, $u_1(\sqrt{w_2^2+p^2}, \sqrt{t_1^2+d^2})$ usw. Integrale dieser Gleichungen, wobei c, e, p, d usw. beliebige Konstanten sein können.

§ 3.

Beziehungen für die neuen Koordinaten und Bemerkungen.

Aus (7) und (8) § 1 und (1) und (2) § 2 folgt

$$(1)\qquad \varkappa r_1 = 2q w_1 t_1;\qquad \varkappa r_2 = 2q w_2 t_2.$$

Ferner haben wir:

$$(2)\quad w_1^2 = \tfrac{1}{2}\cosh\left\{\xi + i\left(\delta+\tfrac{\pi}{2}\right)\right\}+\tfrac{1}{2};\quad t_1^2 = \tfrac{1}{2}\cosh\left\{\xi - i\left(\delta+\tfrac{\pi}{2}\right)\right\}+\tfrac{1}{2}.$$

Demnach ergibt sich wegen (2) § 1

$$(3)\qquad q(w_1^2-t_1^2) = q(w_2^2-t_2^2) = iq\sinh\xi\cdot\cos\delta = i\varkappa x.$$

Aus (1) und (3) fließt ferner

$$(4)\qquad q(w_1+it_1)^2 = i\varkappa r_1(1+\cos\varphi_1) = 2i\varkappa r_1\cos^2\tfrac{\varphi_1}{2}$$

und

$$(5)\qquad q(w_1-it_1)^2 = -2i\varkappa r_1\sin^2\tfrac{\varphi_1}{2},$$

woraus wir erhalten

$$(6)\quad -i\sqrt{iq}(w_1+it_1) = \sqrt{2\varkappa r_1}\cos\tfrac{\varphi_1}{2};\quad -\sqrt{iq}(w_1-it_1) = \sqrt{2\varkappa r_1}\sin\tfrac{\varphi_1}{2};$$

und analog

$$(7)\quad -i\sqrt{iq}\,(w_2+it_2)=\sqrt{2\varkappa r_2}\cos\tfrac{\varphi_2}{2};\quad -\sqrt{iq}\,(w_2-it_2)=\sqrt{2\varkappa r_2}\sin\tfrac{\varphi_2}{2};$$

demnach

$$(8)\qquad w_1=\sqrt{\tfrac{r_1}{2a}}\,e^{\frac{i\varphi_1}{2}+\frac{i\pi}{4}};\qquad t_1=\sqrt{\tfrac{r_1}{2a}}\,e^{\frac{-i\varphi_1}{2}-\frac{i\pi}{4}}$$

und

$$(9)\qquad w_2=\sqrt{\tfrac{r_2}{2a}}\,e^{\frac{i\varphi_2}{2}+\frac{i\pi}{4}};\qquad t_2=\sqrt{\tfrac{r_2}{2a}}\,e^{\frac{-i\varphi_2}{2}-\frac{i\pi}{4}}.$$

Diese Gleichungen (8) und (9) können wir gewissermaßen als Definition für unsere Variablen w_1, t_1, w_2 und t_2 betrachten und sie stimmen mit (1) und (2) § 2 vollkommen überein.

Weiter haben wir

$$(10)\quad q(w_1w_2+t_1t_2)=-q\sinh\xi\cdot\sin\delta=-\varkappa\sqrt{r_1r_2}\sin\tfrac{1}{2}(\varphi_1+\varphi_2),$$

$$(11)\quad q(w_1w_2-t_1t_2)=iq\cosh\xi\cdot\cos\delta=i\varkappa\sqrt{r_1r_2}\cos\tfrac{1}{2}(\varphi_1+\varphi_2)$$

und

$$(12)\qquad q(w_1t_2+w_2t_1)=q\sinh\xi=\varkappa\sqrt{r_1r_2}\cos\tfrac{1}{2}(\varphi_1-\varphi_2),$$

$$(13)\qquad q(w_1t_2-w_2t_1)=-iq\cos\delta=i\varkappa\sqrt{r_1r_2}\cos\tfrac{1}{2}(\varphi_1-\varphi_2).$$

Ferner haben wir

$$(14)\qquad \begin{aligned}\sqrt{q}\,(w_1+t_1)&=\sqrt{2\varkappa r_2}\cos\tfrac{1}{2}\left(\varphi_1+\tfrac{\pi}{2}\right);\\ \sqrt{q}\,(w_1-t_1)&=i\sqrt{2\varkappa r_1}\sin\tfrac{1}{2}\left(\varphi_1+\tfrac{\pi}{2}\right),\end{aligned}$$

woraus die Beziehungen (11) § 2 folgen.

Schreiben wir nun ausführlicher $w_1(\xi,\delta)$ statt w_1 usw., so ergeben sich aus (1) und (2) § 2 noch die Beziehungen

$$(15)\qquad \begin{aligned}w_1(\xi,\pi-\delta)&=-t_1(\xi,\delta);\qquad w_2(\xi,\pi-\delta)=-t_2(\xi,\delta),\\ t_1(\xi,\pi-\delta)&=-w_1(\xi,\delta);\qquad t_2(\xi,\pi-\delta)=-w_2(\xi,\delta)\end{aligned}$$

und

$$(16)\qquad \begin{aligned}w_1(\xi,-\delta)&=it_2(\xi,\delta);\qquad t_1(\xi,-\delta)=-iw_2(\xi;\delta),\\ t_2(\xi,-\delta)&=-iw_1(\xi,\delta);\quad w_2(\xi,-\delta)=it_1(\xi,\delta),\end{aligned}$$

welche wir beim Übergang von dem Punkt P in Fig. 1 zu einem Punkt, welcher zu P symmetrisch in bezug auf die X-Achse, resp. Y-Achse gelegen ist, brauchen werden.

Wir möchten noch einige wichtige Werte unserer Variablen hervorheben. Es wird sein

$$w_1 = t_1 = 0; \quad r_1 = 0; \quad r_2 = 2a; \quad \varphi_2 = \frac{\pi}{2},$$

(17)

$$w_2 = t_2 = 0; \quad r_2 = 0; \quad r_1 = 2a; \quad \varphi_1 = -\frac{\pi}{2},$$

(18) $$w_1 = t_1 = 1; \quad r_2 = 0; \quad r_1 = 2a; \quad \varphi_1 = -\frac{\pi}{2}$$

und

(19) $$w_2 = i; \quad t_2 = -i; \quad r_1 = 0; \quad r_2 = 2a; \quad \varphi_2 = \frac{\pi}{2}.$$

Nun können wir, wie aus der Fig. 1 ersichtlich, zu dem Pol A_2 auch bei dem Wert $\varphi_1 = \frac{3}{2}\pi$ gelangen und analog zu dem Pol A_1 bei dem Wert $\varphi_2 = -\frac{3}{2}\pi$.

Wir erhalten dann

(20) $$w_1 = t_1 = -1; \quad r_2 = 0; \quad r_1 = 2a; \quad \varphi_1 = \frac{3}{2}\pi$$

und

(21) $$w_2 = -i; \quad t_2 = i; \quad r_1 = 0; \quad r_2 = 2a; \quad \varphi_2 = -\frac{3}{2}\pi.$$

Denken wir uns einen unendlich langen Streifen in der YZ-Ebene, welche durch die Z-Achse halbiert wird, so stellt die Linie $A_1 A_2$ in der Fig. 1 den Schnitt dieses Streifens mit der YX-Ebene dar. In diesem Sinn werden wir das Wort Streifen gebrauchen.

Wir haben also auf dem Streifen $\xi = 0$. Außerhalb des Streifens auf der YZ-Ebene resp. auf den entsprechenden Teilen der Y-Achse wird $\delta = \pm\frac{\pi}{2}$ sein.

Laut (8) und (9) erhalten wir demnach auf der positiven ($x > 0$) und negativen Seite des Streifens

(22) $$w_1 = t_1; \quad w_2 = -t_2; \quad x = \xi = 0.$$

Außerhalb des Streifens, auf der positiven Y-Achse wird sein

(23) $$w_1 = -t_1; \quad w_2 = -t_2; \quad x = 0; \quad \delta = \frac{\pi}{2}$$

und auf der negativen Y-Achse

(24) $$w_1 = t_1; \quad w_2 = t_2; \quad x = 0; \quad \delta = -\frac{\pi}{2}.$$

I. Bemerkung zu den Quadratwurzeln in (4) § 2. Setzen wir vorläufig

(a) $$t_2 = \pm\sqrt{t_1^2 - 1},$$

so würde wegen (19) allgemein folgen, daß wir

(b) $$t_2 = -\sqrt{t_1^2 - 1}$$

schreiben müßten.

Wir werden sofort zeigen, daß dies im allgemeinen nicht richtig sein kann. Indem wir z. B.

(c) $$\xi \geqq 0$$

festlegen, ergibt sich aus den zweiten Gleichungen (1) und (2) § 2, für $\delta = \frac{\pi}{2}$, d. h. $\varphi_2 = \frac{\pi}{2}$ (Fig. 1)

(d) $$t_1 = -i \sinh \frac{\xi}{2}; \qquad t_2 = -i \cosh \frac{\xi}{2},$$

was wegen (c) in Übereinstimmung mit (b) sein wird. Machen wir nun aber $\delta = -\frac{\pi}{2}$, d. h. $\varphi_2 = -\frac{\pi}{2}$, so folgt

(e) $$t_1 = \cosh \frac{\xi}{2}; \qquad t_2 = \sinh \frac{\xi}{2},$$

was wieder wegen (c) dem Ausdruck (b) widerspricht. Statt letzteren müßten wir in diesem Fall schreiben:

(f) $$t_2 = \sqrt{t_1^2 - 1}\,.$$

Dieser Widerspruch erklärt sich dadurch, daß die Amplitude von t_2, wie aus (9) ersichtlich, bei dem Durchgang durch den Pol A_2 längs der Y-Achse einen Sprung von $\frac{i\pi}{2}$ macht, während der Betrag von t_2 dabei stetig durch Null geht.

Analoge Verhältnisse treffen wir auch bei den anderen Wurzeln in (4) § 2.

Um alle Zweideutigkeit zu vermeiden, werden wir die Ausdrücke (4) § 2 wie sie hingeschrieben sind gebrauchen, dabei aber beachten, daß *nicht* die Quadratwurzeln die linken Seiten bestimmen, sondern umgekehrt, es sollen *die linken Seiten*, welche durch (8) und (9) festgelegt sind, die Wurzeln *bestimmen*, d. h. überall, wo solche Wurzeln vorkommen, müssen sie durch die *linken Seiten von* (4) § 2 *ersetzt werden*, und dann erst soll mit diesen Größen gerechnet werden.

An dieser Regel wollen wir im folgenden immer festhalten.

II. Bemerkung zu dem Übergang zu $a = 0$. Es wird sein:

(g) $$\sqrt{q}\, t_1 = \sqrt{\frac{\varkappa r_1}{2}}\, e^{\frac{-i\varphi_1}{2} - \frac{i\pi}{4}}; \qquad \sqrt{q}\, t_2 = \sqrt{\frac{\varkappa r_2}{2}}\, e^{\frac{-i\varphi_2}{2} - \frac{i\pi}{4}}.$$

Wir wollen nun den Pol A_1 in der Fig. 1 festhalten und dabei a zu Null konvergieren lassen. Es wird dann, wie aus der Figur ersichtlich,

$r_1 = r_2$ und $\varphi_1 = \varphi_2$ und die beiden Größen in (g) werden einander gleich. Ähnliche Verhältnisse gelten auch für die anderen Koordinaten.

Haben wir nun einen Ausdruck mit unseren Variablen so umgeformt, daß dieselben in der Form (g) erscheinen, so können wir dann den Übergang zu $a = 0$ ausführen, wobei das q, welches außerdem noch allein auftritt, verschwinden muß.

Nachdem dieser Übergang stattgefunden hat, können wir wieder zu unseren Variablen und zu dem q zurückkehren, d. h. nur mit *einem* Pol A_1 operieren, wobei jetzt das q nur als Begleitstück unserer Koordinaten auftritt.

Multiplizieren wir (b) und (f) beiderseits mit q, so erhalten wir

(h) $$\sqrt{q}\,t_2 = -\sqrt{q\,t_1^2 - q}; \qquad \sqrt{q}\,t_2 = \sqrt{q\,t_1^2 - q}.$$

Bei dem Übergang zu $a = 0$ erhalten wir, laut dem eben Gesagten, aus der ersten Beziehung $\sqrt{q}\,t_1 = -\sqrt{q}\,t_1$ und aus der zweiten $\sqrt{q}\,t_1 = \sqrt{q}\,t_1$. Dieses bestätigt noch einmal die Richtigkeit der Bemerkung I.

Der Übergang zu $a = 0$ wird später an Beispielen noch klarer werden. (Siehe § 11 und 12.)

III. Bemerkung zu dem Übergang zu $a = \infty$. Auch jetzt halten wir den Pol A_1 fest und rücken nun mit dem Pol A_2 nach unten ins Unendliche. In diesem Sinn werden wir den Übergang zu $a = \infty$ verstehen. Es ändert sich hierbei die erste Größe in (g) nicht, während die zweite unendlich wird. Aus der Fig. 1 ist ersichtlich, daß bei endlichem r_1, φ_2 zu dem Werte $\frac{\pi}{2}$ strebt und r_2 zu $r_2 = 2a = \infty$.

Bei diesem Übergang eines entsprechenden Ausdruckes (wo die Variablen wieder in der Form (g) geschrieben sind), müssen die Größen $\sqrt{q}\,t_2$ und $\sqrt{q}\,w_2$ verschwinden und ebenfalls die freien q.

Nach vollzogenem Übergang wird der Ausdruck nur die Größen $\sqrt{q}\,t_1$ und $\sqrt{q}\,w_1$ enthalten können, und wir dürfen dann, analog wie in der Bemerkung II, wieder mit den Variablen w_1 und t_1 und mit der Größe q operieren. (Siehe § 11 und 12.)

§ 4.

Verschiedene Typen von Integralen der Wellengleichung.

Als einfachste Typen von Integralen der Gleichungen (7) bis (9) § 2 erhalten wir

(1) $$\frac{J_\nu(2q\,w_1\,t_1)\,t_1^{2\nu}}{(w_1\,t_1)^\nu}; \qquad \frac{J_\nu(2q\,w_1\,t_1)\,w_1^{2\nu}}{(w_1\,t_1)^\nu},$$

(2) $$\frac{J_\nu(2q\,w_1\,t_2)\,t_2^{2\nu}}{(w_1\,t_2)^\nu}; \qquad \frac{J_\nu(2q\,w_1\,t_2)\,w_1^{2\nu}}{(w_1\,t_2)^\nu}$$

und

$$(3) \qquad \frac{J_\nu(2q w_2 t_1)\, t_1^{2\nu}}{(w_2 t_1)^\nu}, \qquad \frac{J_\nu(2q w_2 t_1)\, w_2^{2\nu}}{(w_2 t_1)^\nu},$$

wobei ν beliebig sein kann.

Auf Grund von § 3 können wir statt (1) auch schreiben

$$(4) \qquad J_\nu(\varkappa r_1)\, e^{-i\nu\varphi_1 - \frac{i\nu\pi}{2}}; \qquad J_\nu(\varkappa r_1)\, e^{+i\nu\varphi_1 + \frac{i\nu\pi}{2}},$$

was bekannte Integrale darstellen und deshalb kein besonderes Interesse bieten.

Mehr von Interesse sind die Integrale (2) und (3), da dort beide Pole zugleich vorkommen.

Wesentlicher sind die Integrale, die man auf Grund der Bemerkung am Schlusse des § 2 erhalten kann. Z. B.

$$(5) \qquad \frac{J_\nu\left(2q\sqrt{w_1^2+c^2}\sqrt{t_2^2+d^2}\right)\left(\sqrt{t_2^2+d^2}\right)^{2\nu}}{\left(\sqrt{w_1^2+c^2}\sqrt{t_2^2+d^2}\right)^\nu},$$

wobei w_2 statt w_1 und t_1 statt t_2 stehen kann.

Hieraus erhalten wir sofort folgende Typen von Integralen

$$(6) \qquad \int_b^m f(\lambda)\, \frac{J_\nu\left(2q\sqrt{w_1^2+c^2}\sqrt{t_2^2-\lambda^2}\right)\left(\sqrt{t_2^2-\lambda^2}\right)^{2\nu} d\lambda}{\left(\sqrt{w_1^2+c^2}\sqrt{t_2^2-\lambda^2}\right)^\nu}$$

oder

$$(7) \qquad \int_b^m f(\lambda)\, \frac{J_\nu\left(2q\sqrt{w_1^2+\lambda^2}\sqrt{t_2^2-\lambda^2}\right)\left(\sqrt{t_2^2-\lambda^2}\right)^{2\nu} d\lambda}{\left(\sqrt{w_1^2+\lambda^2}\sqrt{t_1^2-\lambda^2}\right)^\nu}$$

oder auch

$$(8) \qquad \int_b^m d\lambda \int_c^d f(\lambda, z)\, \frac{J_\nu\left(2q\sqrt{w_1^2+z^2}\sqrt{t_2^2-\lambda^2}\right)\left(\sqrt{t_2^2-\lambda^2}\right)^{2\nu} dz}{\left(\sqrt{w_1^2+z^2}\sqrt{t_2^2-\lambda^2}\right)^\nu}.$$

Noch wichtiger sind aber die Integrale vom Typus

$$(9) \qquad \int_b^{w_2} f(\lambda)\, \frac{J_\nu\left(2q\sqrt{t_1^2+c^2}\sqrt{w_2^2-\lambda^2}\right)\left(\sqrt{w_2^2-\lambda^2}\right)^{2\nu} d\lambda}{\left(\sqrt{t_1^2+c^2}\sqrt{w_2^2-\lambda^2}\right)^\nu}$$

oder

$$(10) \qquad \int_b^{w_2} f(\lambda)\, \frac{J_\nu\left(2q\sqrt{t_1^2+\lambda^2}\sqrt{w_2^2-\lambda^2}\right)\left(\sqrt{w_2^2-\lambda^2}\right)^{2\nu} d\lambda}{\left(\sqrt{t_1^2+\lambda^2}\sqrt{w_2^2-\lambda^2}\right)^\nu},$$

wobei $\nu \geqq 0$ sein muß.

Während bei den Integralen (6) bis (8) unter den Wurzeln das Vorzeichen vor c^2, λ^2, z^2 beliebig sein kann, muß aber in (9) resp. (10) das Vorzeichen vor λ^2 in derjenigen Wurzel, welche der *oberen Grenze entspricht*, negativ sein. Demnach können wir in (10) auch $\sqrt{t_1^2-\lambda^2}$ schreiben.

Für das Folgende sind vorläufig von besonderer Wichtigkeit Integrale vom Typus

$$\int_b^{w_2} f(\lambda)\, J_0(2q\sqrt{t_1^2+\lambda^2}\,\sqrt{w_2^2-\lambda^2})\, d\lambda, \tag{11}$$

wobei also b eine Konstante bedeutet und $f(\lambda)$ eine willkürliche Funktion von λ. Verschiedene Arten solcher Integrale (11) erhält man durch Vertauschung von w_1 und t_1 resp. w_2 und t_2 miteinander, unter Beachtung der Beziehungen (3) § 2.

Wir betrachten nun das bekannte Sommerfeldsche Integral und erhalten, unter Beachtung der Beziehungen (3) und (6) § 3,

$$A = e^{-i\varkappa x} \int_0^{\sqrt{2\varkappa r_1}\sin\frac{\varphi_1}{2}} e^{-i\lambda^2}\, d\lambda = -e^{q(t_1^2-w_1^2)} \int_0^{\sqrt{iq}\,(w_1-it_1)} e^{-i\lambda^2}\, d\lambda, \tag{12}$$

wobei also die rechte Seite ein Integral der Gleichung (7) § 2 sein muß, wovon man sich auch leicht überzeugen kann. Nun können wir auch schreiben

$$A = -\sqrt{iq}\, e^{-qw_1^2} \int_0^{w_1} e^{q\lambda^2} \cos(2qt_1\lambda)\, d\lambda + i\sqrt{iq}\, e^{qt_1^2} \int_0^{t_1} e^{-q\lambda^2} \cos(2qw_1\lambda)\, p\lambda, \tag{13}$$

und hierbei stellt rechterseits jedes Glied *einzeln* ein Integral der Gleichung (7) § 2 dar.

Dadurch sind wir wieder zu neuen Typen von Integralen gekommen, wenigstens der Form nach, denn wie wir in § 11 sehen werden, lassen sich die Integrale (13) durch Integrale vom Typus (11) darstellen.

Die Integrale (13) sind zweiwertig in bezug auf den Pol A_1.

Sie lassen sich erweitern und wir kommen zu dem Typus

$$e^{-qw_1^2} \int_0^{w_1} e^{q\lambda^2} f(w_2^2-\lambda^2) \cos(2qt_1\lambda)\, d\lambda, \tag{14}$$

wobei f eine beliebige Funktion des Argumentes $w_2^2-\lambda^2$ sein kann.

Zum Schluß führen wir noch ein Integral an, und zwar

$$B = e^{-i\varkappa x} \int_0^{\sqrt{2\varkappa r_1}\sin^1/_2\varphi_1} e^{-i\lambda^2} d\lambda \cdot \int_0^{\sqrt{2\varkappa r_1}\cos^1/_2\varphi_1} e^{+i\lambda^2} d\lambda \tag{15}$$

$$= -qe^{q(t_1^2-w_1^2)} \int_0^{w_1-it_1} e^{q\lambda^2} d\lambda \cdot \int_0^{w_1+it_1} e^{q\lambda^2} d\lambda,$$

wofür wir auch schreiben können

$$B = -qe^{-q(t_1^2+w_1^2)} \left\{ \int_0^{w_1} e^{q\lambda^2} \cos(2qt_1\lambda)\, d\lambda \right\}^2 \tag{16}$$

$$- qe^{q(t_1^2+w_1^2)} \left\{ \int_0^{t_1} e^{-q\lambda^2} \cos(2qw_1\lambda)\, d\lambda \right\}^2.$$

Dieses Integral ist augenscheinlich einwertig.

§ 5.

Funktionentheoretische Betrachtungen.

Wir führen jetzt folgende Funktionen ein:

1) $$M = \int_0^{w_2} \sqrt{\lambda^2 + 1}\, f(\lambda^2, w_2^2, t_2^2)\, d\lambda$$

und

2) $$N = \int_1^{w_1} \sqrt{\lambda^2 - 1}\, \psi(\lambda^2, w_2^2, t_2^2)\, d\lambda .$$

Vertauschen wir in einer Funktion w_1 und t_1 resp. w_2 und t_2 miteinander, so werden wir die so entstandene neue Funktion durch einen Strich bezeichnen. Bei diesem Tausch müssen wir immer (3) § 2 in Betracht ziehen.

Aus (1) erhalten wir demnach

3) $$M' = \int_0^{t_2} \sqrt{\lambda^2 + 1}\, f(\lambda^2, t_2^2, w_2^2)\, d\lambda$$

und aus (2)

4) $$N' = \int_1^{t_1} \sqrt{\lambda^2 - 1}\, \psi(\lambda^2, t_1^2, w_1^2)\, d\lambda .$$

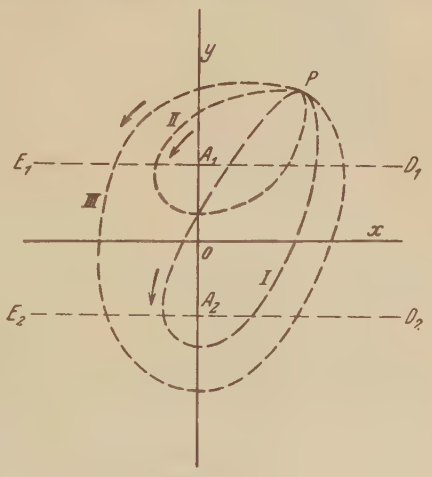


Fig. 3.

Bezüglich f und ψ wollen wir festlegen, daß sie *ganze* Funktionen von λ^2, w_1^2, t_1^2 bzw. w_2^2, t_2^2 sein sollen.

Wir denken uns jetzt M für den Punkt P der Fig. 1 gegeben und wollen nun untersuchen, wie sich M verhält nach dem Umlauf, im Sinne der Pfeile, längs den verschiedenen Schleifen I, II, III der Fig. 3 in der *reellen* λX-Ebene.

Den Wert der Funktion nach Umlauf einer Schleife bezeichnen wir durch einen Index, welcher der Schleife entspricht.

Zwecks Untersuchung der hier auftretenden Verhältnisse müssen wir die Ebene der komplexen Veränderlichen λ in Betracht ziehen (Fig. 4). Auf dieser Ebene bilden wir zugleich die Variable w_2 ab. Da wir $\lambda = \varrho e^{i\vartheta}$ schreiben können, so erhalten wir demnach

5) $$w_2 = \sqrt{\frac{r_2}{2a}}\, e^{\frac{i\varphi_2}{2} + \frac{i\pi}{4}} = \varrho e^{i\vartheta}; \qquad \varrho = \sqrt{\frac{r_2}{2a}}; \qquad \vartheta = \frac{\varphi_2}{2} + \frac{\pi}{4}.$$

D. h. für $\varphi_2 = 0$ ist $\vartheta = \frac{\pi}{4}$. Ziehen wir deshalb durch O, Fig. 4, eine

Gerade COD unter 45^0 zur X-Achse, so entspricht OD der Fig. 4 der Geraden A_2D_2 der Fig. 3 und O dem Pole A_2. Die Halbebene oberhalb CD (in Richtung von OE) entspricht der gesamten Ebene der Fig. 3. Endlich entsprechen die Geraden OY und OE der Fig. 4 den Geraden A_2Y und A_2E_2 der Fig. 3. Es stellt demnach die Fig. 3 das eine (obere) Blatt der Funktion w_2 dar, wobei der Verzweigungsschnitt vom Pol A_2 längs A_2D_2 verläuft. Die Gerade OC der Fig. 4 entspricht dem anderen Ufer dieses Verzweigungsschnittes.

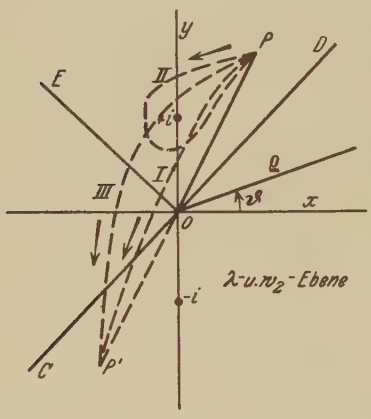


Fig. 4.

Es liegt demnach der Punkt P der Fig. 3, welcher einem bestimmten Wert von w_2 im oberen Blatt entpricht, in der Fig. 4 zwischen OD und OY und wir markieren ihn dort ebenfalls durch P.

Infolge der Umlaufsrichtung in der Fig. 3 ersehen wir, daß der Pol A_1 dem Werte $\varphi_2 = \frac{\pi}{2}$ der Größe w_2 entsprechen wird. Laut (19) § 3 ist hierbei $w_2 = i$. Also entspricht der Punkt i der Fig. 4 dem Pole A_1. Zugleich ist aber i ein Verzweigungspunkt der Wurzel in (1). Wir werden später sehen, daß hierin das Wesentliche der Sache enthalten ist.

Aus dem eben Gesagten erhellt, daß die Strecke Oi der Fig. 4 der Strecke A_1A_2 zwischen den Polen der Fig. 3 entsprechen wird.

Nun können wir diejenigen Wege in der Fig. 4 bestimmen, welche den Schleifen der Fig. 3 entsprechen werden.

Zu dem Zweck zeichnen wir den Punkt P' (Fig. 4), welcher das Spiegelbild des Punktes P in bezug auf den Koordinatenanfang bildet. Diesem Punkt P' entspricht der Wert $-w_2$ und liegt, wie aus der Fig. 4 ersichtlich, im unteren Blatt der Fig. 3.

Betrachten wir nun jetzt die Schleife I, so ersehen wir erstens, daß wir zu dem Wert $-w_2$ gelangen, und zweitens, daß wir dabei die Y-Achse in der Fig. 3 *zwischen* den beiden Polen schneiden. Demnach entspricht der Schleife I der Fig. 3 der Weg I der Fig. 4 in Richtung des Pfeiles.

Bei der Schleife II ersehen wir, da wir den Pol A_2 nicht umkreisen, daß wir zu dem Anfangswert von w_2 zurückkehren und daß wir die Y-Achse *zuerst oberhalb* des Poles A_1 schneiden. Demnach entspricht die Schleife II der Schleife II in der Fig. 4 um den Punkt i herum in Richtung des Pfeiles.

Endlich haben wir bei der Schleife III einen Umlauf um beide Pole A_1 und A_2, wobei wir zu dem Wert $-w_2$ kommen und die Y-Achse *zuerst*

oberhalb des Poles A_1 geschnitten wird. Es entspricht demnach der Schleife III der Weg III in der Fig. 4 in Richtung des Pfeiles.

Der Anfangswert (1) von M wird einfach durch das Integral von O bis P in der λ-Ebene der Fig. 4 festgelegt.

Demnach erhalten wir die Werte von M_{I}, M_{II} und M_{III} einfach durch Addition zu diesem Anfangswert M der Integrale längs der Wege I, II und III der Fig. 4. Wegen der Eigenschaft der Funktion f, da wir dabei immer zu dem Anfangswert z. B. von w_2 resp. $-w_2$ kommen, spielen hierbei w und t in f keine Rolle.

Deshalb erhalten wir für den Wert von M_{I}, indem wir den Weg I deformieren von P bis O und von O bis P',

$$M_{\mathrm{I}} = M + \int_{w_2}^{-w_2} = M + \int_{0}^{-w_2} - \int_{0}^{w_2} = -M, \tag{6}$$

was übrigens auch direkt aus (1) folgt.

Analog erhalten wir

$$M_{\mathrm{II}} = M + \int_{w_2}^{0} + \int_{0}^{i} - \int_{i}^{0} - \int_{0}^{w_2} \tag{7}$$

oder

$$M_{\mathrm{II}} = -M + 2i \int_{0}^{1} \sqrt{1-\lambda^2}\, f(-\lambda^2, w_2^2, t_2^2)\, d\lambda \tag{8}$$

und

$$M_{\mathrm{III}} = M + \int_{w_2}^{0} + \int_{0}^{i} - \int_{i}^{0} - \int_{0}^{-w_2}, \tag{9}$$

d. h.

$$M_{\mathrm{III}} = M + 2i \int_{0}^{1} \sqrt{1-\lambda^2}\, f(-\lambda^2, w_2^2, t_2^2)\, d\lambda. \tag{10}$$

Wir wenden uns jetzt zu der Funktion (4) und wollen die entsprechenden Werte N_{I}', N_{II}' und N_{III}' bestimmen. Zu dem Zweck bilden wir ebenfalls t_1 auf die λ-Ebene ab (Fig. 5). Es wird sein

$$t_1 = \sqrt{\frac{r_1}{2a}}\, e^{-\frac{i\varphi_1}{2} - \frac{i\pi}{4}} = \varrho\, e^{i\vartheta}; \tag{11}$$

$$\varrho = \sqrt{\frac{r_1}{2a}}; \qquad \vartheta = -\frac{\varphi_1}{2} - \frac{\pi}{4}.$$

Ziehen wir nun durch O, Fig. 5, eine Gerade unter -45^0 zur X-Achse, so entspricht die Gerade OD der Fig. 5 der Geraden $A_1 D_1$ der Fig. 3 und O dem

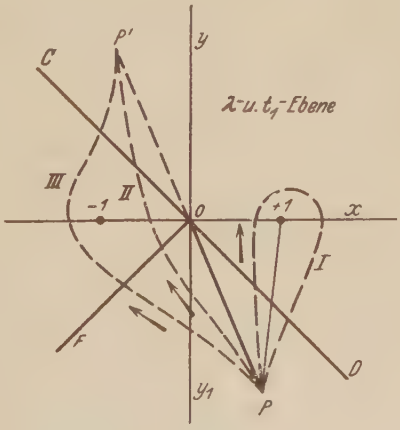


Fig. 5.

Pole A_1. Die Halbebene unterhalb CD (in Richtung von OE) entspricht der gesamten Ebene der Fig. 3. Die Geraden OY_1 und OE der Fig. 5 entsprechen den Geraden A_1Y und A_1E_1 der Fig. 3.

Die negative X-Achse der Fig. 5 entspricht der Y-Achse der Fig. 3 auch in negativer Richtung, vom Pole A_1 angefangen.

Unter Beachtung, daß beim Umlauf um I die Wurzel in (4) ihr Zeichen ändert, erhalten wir nun

$$N_{\mathrm{I}}' = N' + \int_{t_1}^{1} - \int_{1}^{t_1} \tag{12}$$

oder

$$N_{\mathrm{I}}' = -N'; \tag{13}$$

$$N_{\mathrm{II}}' = N' + \int_{t_1}^{1} + \int_{1}^{0} + \int_{0}^{-t_1} \tag{14}$$

oder

$$N_{\mathrm{II}}' = -N' - 2\int_{0}^{1} \sqrt{\lambda^2 - 1}\, \psi(\lambda^2, t_1^2, w_1^2)\, d\lambda. \tag{15}$$

Hierbei ist die Integration längs der reellen Achse zu führen, wobei also ständig $\lambda^2 \leqq 1$ sein wird. Wir kommen nun nicht weiter, ohne eine Voraussetzung über den Wurzelwert zu machen. Diese Voraussetzung muß im Einklang sein mit der Bemerkung I § 3 und demgemäß mit den späteren analytischen Betrachtungen. Wir setzen deshalb fest, daß für das Integral in (15) es $\sqrt{1-\lambda^2} = i\sqrt{\lambda^2 - 1}$ gelten soll, d. h.

$$\sqrt{\lambda^2 - 1} = -i\sqrt{\lambda^2 - 1}, \tag{16}$$

und erhalten demnach statt (15)

$$N_{\mathrm{II}}' = -N' + 2i\int_{0}^{1} \sqrt{1-\lambda^2}\, \psi(\lambda^2, t_1^2, w_1^2)\, d\lambda. \tag{17}$$

Endlich ist

$$N_{\mathrm{III}}' = N' + 2i\int_{0}^{1} \sqrt{1-\lambda^2}\, \psi(\lambda^2, t_1^2, w_1^2)\, d\lambda. \tag{18}$$

Wir wenden uns jetzt zu den Integralen M' in (3) und N in (2) und erhalten

$$M_{\mathrm{I}}' = -M', \tag{19}$$

$$M_{\mathrm{II}}' = -M' - 2i\int_{0}^{1} \sqrt{1-\lambda^2}\, f(-\lambda^2, t_2^2, w_2^2)\, d\lambda, \tag{20}$$

$$M_{\mathrm{III}}' = M' - 2i\int_{0}^{1} \sqrt{1-\lambda^2}\, f(-\lambda^2, t_2^2, w_2^2)\, d\lambda, \tag{21}$$

$$N_{\mathrm{I}} = -N, \tag{22}$$

$$N_{\mathrm{II}} = -N - 2i\int_0^1 \sqrt{1-\lambda^2}\,\psi(\lambda^2, w_1^2, t_1^2)\,d\lambda, \tag{23}$$

$$N_{\mathrm{III}} = N - 2i\int_0^1 \sqrt{1-\lambda^2}\,\psi(\lambda^2, w_1^2, t_1^2)\,d\lambda, \tag{24}$$

wobei wir hier statt (16) die Beziehung

$$\sqrt{\lambda^2-1} = i\sqrt{1-\lambda^2} \tag{25}$$

gebrauchen müssen (siehe § 8).

Nun führen wir folgende Integrale ein:

$$R = \int_0^{w_2} \frac{\chi(\lambda^2, w_2^2, t_2^2)\,d\lambda}{\sqrt{\lambda^2+1}}, \tag{26}$$

$$W = \int_1^{w_1} \frac{\gamma(\lambda^2, w_1^2, t_1^2)\,d\lambda}{\sqrt{\lambda^2-1}}. \tag{27}$$

Vertauschen wir hier w und t miteinander, so folgt

$$R' = \int_0^{t_2} \frac{\chi(\lambda^2, t_2^2, w_2^2)\,d\lambda}{\sqrt{\lambda^2+1}} \tag{28}$$

und

$$W' = \int_1^{t_1} \frac{\gamma(\lambda^2, t_1^2, w_1^2)\,d\lambda}{\sqrt{\lambda^2-1}}. \tag{29}$$

In Analogie mit dem Vorhergehenden erhalten wir hieraus, unter Berücksichtigung von (16) resp. (25),

$$R_{\mathrm{I}} = -R, \tag{30}$$

$$R_{\mathrm{II}} = -R + 2i\int_0^1 \frac{\chi(-\lambda^2, w_2^2, t_2^2)\,d\lambda}{\sqrt{1-\lambda^2}}, \tag{31}$$

$$R_{\mathrm{III}} = R + 2i\int_0^1 \frac{\chi(-\lambda^2, w_2^2, t_2^2)\,d\lambda}{\sqrt{1-\lambda^2}} \tag{32}$$

und

$$W_{\mathrm{I}}' = -W', \tag{33}$$

$$W_{\mathrm{II}}' = -W' - 2i\int_0^1 \frac{\gamma(\lambda^2, t_1^2, w_1^2)\,d\lambda}{\sqrt{1-\lambda^2}}, \tag{34}$$

$$W_{\mathrm{III}}' = -W' - 2i\int_0^1 \frac{\gamma(\lambda^2, t_1^2, w_1^2)\,d\lambda}{\sqrt{1-\lambda^2}}. \tag{35}$$

Ferner folgt

$$R_{\mathrm{I}}' = -R', \tag{36}$$

$$R_{\mathrm{II}}' = -R' - 2i\int_0^1 \frac{\chi(-\lambda^2, t_2^2, w_2^2)\,d\lambda}{\sqrt{1-\lambda^2}}, \tag{37}$$

$$R_{\mathrm{III}}' = R - 2i\int_0^1 \frac{\chi(-\lambda^2, t_2^2, w_2^2)\,d\lambda}{\sqrt{1-\lambda^2}} \tag{38}$$

und

$$W_{\mathrm{I}} = -W, \tag{39}$$

$$W_{\mathrm{II}} = -W + 2i\int_0^1 \frac{\gamma(\lambda^2, w_1^2, t_1^2)\,d\lambda}{\sqrt{1-\lambda^2}}, \tag{40}$$

$$W_{\mathrm{III}} = W + 2i\int_0^1 \frac{\gamma(\lambda^2, w_1^2, t_1^2)\,d\lambda}{\sqrt{1-\lambda^2}}. \tag{41}$$

Es ist hierbei als selbstverständlich angenommen, daß die Funktionen χ und γ von derselben Art wie f und ψ sind.

§ 6.

Doppelpolige Funktionen.

Unter einer doppelpoligen Funktion verstehen wir eine solche, die den folgenden Bedingungen genügt:

I. Sie ist einwertig in bezug auf die Schleife III (Fig. 3),

II. zweiwertig in bezug auf die Schleife I resp. II, und

III. endlich in den Polen A_1 und A_2.

Wir gehen nun jetzt zur Bildung solcher Funktionen über. Zu dem Zweck führen wir die folgenden Bedingungen ein:

(1) $f(-\lambda^2, w_2^2, t_2^2) = \psi(\lambda^2, t_1^2, w_1^2)$ oder $f(-\lambda^2, t_2^2, w_2^2) = \psi(\lambda^2, w_1^2, t_1^2)$

und

(2) $\chi(-\lambda^2, w_2^2, t_2^2) = \gamma(\lambda^2, t_1^2, w_1^2)$ oder $\chi(-\lambda^2, t_2^2, w_2^2) = \gamma(\lambda^2, w_1^2, t_1^2)$.

Auf Grund des vorigen Paragraphen ersehen wir nun leicht, daß die vier Funktionen

(1) $$u_1 = M - N'; \quad u_2 = M' - N$$

und

(2) $$u_3 = R + W'; \quad u_4 = R' + W$$

im obigen Sinn doppelpolig sein werden.

Im nächsten Paragraphen werden wir sehen, daß diese Funktionen nicht nur doppelpolig, sondern — wie für die Lösung notwendig — auch symmetrisch in bezug auf die X-Achse sein werden, falls wir noch gewisse Bedingungen bezüglich f, ψ, χ und γ einführen.

§ 7.

Analytische Untersuchungen der Funktionen des § 5.

Wir haben folgende unbestimmte Integrale

1) $$\int x^{2m}\sqrt{x^2+1}\,dx = -\frac{(-1)^m x(x^2+1)^{3/2}\Gamma\left(m+\frac{1}{2}\right)}{2\Gamma(m+2)}\sum_{n=0}^{n=m-1}\frac{(-1)^n\Gamma(n+2)x^{2n}}{\Gamma\left(n+\frac{3}{2}\right)} + \frac{(-1)^m\Gamma\left(m+\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma(m+2)}\int\sqrt{x^2+1}\,dx$$

und

2) $$\int\frac{x^{2m}}{\sqrt{x^2+1}}dx = -x^{2m+1}\sqrt{x^2+1}+2(m+1)\int x^{2m}\sqrt{x^2+1}\,dx.$$

Aus (1) und (2) folgt

3) $$\int\frac{x^{2m}dx}{\sqrt{x^2+1}} = -(-1)^m\frac{x(x^2+1)^{3/2}\Gamma\left(m+\frac{1}{2}\right)}{\Gamma(m+1)}\sum_{n=0}^{n=m-1}\frac{(-1)^n\Gamma(n+2)x^{2n}}{\Gamma\left(n+\frac{3}{2}\right)} - x^{2m+1}\sqrt{x^2+1}+\frac{2(-1)^m\Gamma\left(m+\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma(m+1)}\int\sqrt{x^2+1}\,dx.$$

Ferner ist

4) $$\int x^{2m}\sqrt{x^2-1}\,dx = \frac{x(x^2-1)^{3/2}\Gamma\left(m+\frac{1}{2}\right)}{2\Gamma(m+2)}\sum_{n=0}^{n=m-1}\frac{\Gamma(n+2)x^{2n}}{\Gamma\left(n+\frac{3}{2}\right)} + \frac{\Gamma\left(m+\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma(m+2)}\int\sqrt{x^2-1}\,dx$$

und

5) $$\int\frac{x^{2m}dx}{\sqrt{x^2-1}} = x^{2m+1}\sqrt{x^2-1}-2(m+1)\int x^{2m}\sqrt{x^2-1}\,dx,$$

weshalb sich wegen (4) ergibt

6) $$\int\frac{x^{2m}dx}{\sqrt{x^2-1}} = -\frac{x(x^2-1)^{3/2}\Gamma\left(m+\frac{1}{2}\right)}{\Gamma(m+1)}\sum_{n=0}^{n=m-1}\frac{\Gamma(n+2)x^{2n}}{\Gamma\left(n+\frac{3}{2}\right)} + x^{2m+1}\sqrt{x^2-1}-\frac{2\Gamma\left(m+\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma(m+1)}\int\sqrt{x^2-1}\,dx.$$

Natürlich hat man noch

$$\int \sqrt{x^2+1}\,dx = \frac{1}{2}x\sqrt{x^2+1} + \frac{1}{2}\lg(x+\sqrt{x^2+1}) \tag{7}$$

und

$$\int \sqrt{x^2-1}\,dx = \frac{1}{2}x\sqrt{x^2-1} - \frac{1}{2}\lg(x+\sqrt{x^2-1}). \tag{8}$$

Auf Grund der Eigenschaften der Funktionen f und ψ des § 5 können wir schreiben

$$f(\lambda^2, w_2^2, t_2^2) = \sum_{s=0}^{\infty} A_{2s}(w_2^2, t_2^2)\,\lambda^{2s} \tag{9}$$

und

$$\psi(\lambda^2, w_1^2, t_1^2) = \sum_{s=0}^{\infty} B_{2s}(w_1^2, t_1^2)\,\lambda^{2s}. \tag{10}$$

Wir erhalten demnach aus (1) § 5 auf Grund von (1) und (9)

$$M = -\frac{w_2 w_1^3}{2}\sum_{s=0}^{\infty} A_{2s}(-1)^s \frac{\Gamma\left(s+\frac{1}{2}\right)}{\Gamma(s+2)} \sum_{n=0}^{n=s-1} (-1)^n \frac{\Gamma(n+2)\,w_2^{2n}}{\Gamma\left(n+\frac{3}{2}\right)} + \frac{1}{\Gamma\left(\frac{1}{2}\right)}\sum_{s=0}^{\infty}(-1)^s A_{2s}\frac{\Gamma\left(s+\frac{1}{2}\right)}{\Gamma(s+2)} \cdot \int_0^{w_2}\sqrt{\lambda^2+1}\,d\lambda. \tag{11}$$

Berücksichtigen wir nun, daß

$$\frac{\Gamma\left(s+\frac{1}{2}\right)}{\Gamma(s+2)} = \frac{4}{\Gamma\left(\frac{1}{2}\right)}\int_0^1 \lambda^{2s}\sqrt{1-\lambda^2}\,d\lambda \tag{12}$$

sein wird, so können wir statt (11) schreiben

$$M = w_2 w_1^3 S(w_2^2, t_2^2) + \frac{2}{\pi}\{w_1 w_2 + \lg(w_1+w_2)\}\int_0^1 \sqrt{1-\lambda^2}\,f(-\lambda^2, w_2^2, t_2^2)\,d \tag{13}$$

wobei

$$S(w_2^2, t_2^2) = \sum_{s=0}^{\infty} A_{2s}(w_2^2, t_2^2)\cdot\sigma_s(w_2^2) \tag{14}$$

und

$$\sigma_s(w_2^2) = -\frac{\Gamma\left(s+\frac{1}{2}\right)(-1)^s}{\Gamma(s+2)}\sum_{n=0}^{n=s-1}\frac{(-1)^n\,\Gamma(n+2)\,w_2^{2n}}{\Gamma\left(n+\frac{3}{2}\right)} \tag{15}$$

bedeuten.

Auf analoge Weise erhalten wir aus (2) § 5 auf Grund von (4) und (10)

$$N = w_1 w_2^3 T(w_1^2, t_1^2) + \frac{2}{\pi}\{w_1 w_2 - \lg(w_1+w_2)\}\int_0^1 \sqrt{1-\lambda^2}\,\psi(\lambda^2, w_1^2, t_1^2)\,d \tag{16}$$

wobei

$$T(w_1^2, t_1^2) = \sum_{s=0}^{\infty} B_{2s}(w_1^2, t_1^2)\, \tau_s(w_1^2) \tag{17}$$

und

$$\tau_s(w_1^2) = \frac{\Gamma\left(s+\frac{1}{2}\right)}{\Gamma(s+2)} \sum_{n=0}^{n=s-1} \frac{\Gamma(n+2)\, w_1^{2n}}{\Gamma\left(n+\frac{3}{2}\right)} \tag{18}$$

bedeuten.

Ein Vergleich von (15) und (18) ergibt uns

$$\tau_s(x^2) = -(-1)^s \sigma_s(-x^2) \quad \text{oder} \quad \sigma(x^2) = -(-1)^s \tau_s(-x^2). \tag{19}$$

Gehen wir in einer Funktion von dem Punkt P (Fig. 1 oder 3) zu einem Punkt P_1 über, welcher in bezug zu P symmetrisch zur X-Achse liegt, so werden wir die hierbei veränderte Funktion durch einen Querstrich bezeichnen.

Wir erhalten demnach aus (13) unter Berücksichtigung von (16) § 3

$$\overline{M} = t_1 t_2^3\, S(-t_1^2, -w_1^2) + \frac{2}{\pi}\left\{\frac{i\pi}{2} - t_1 t_2 + \lg(t_1+t_2)\right\} \int_0^1 \sqrt{1-\lambda^2}\, f(-\lambda^2, -t_1^2, -w_1^2)\, d\lambda. \tag{20}$$

Wir vertauschen nun in (16) w und t miteinander, d. h. gehen zu der Funktion N' über und erhalten

$$N' = t_1 t_2^3\, T(t_1^2, w_1^2) + \frac{2}{\pi}\{t_1 t_2 - \lg(t_1+t_2)\} \int_0^1 \sqrt{1-\lambda^2}\, \psi(\lambda^2, t_1^2, w_1^2)\, d\lambda. \tag{21}$$

Jetzt führen wir folgende Bedingung ein

$$f(-\lambda^2, -t_1^2, -w_1^2) = \psi(\lambda^2, t_1^2, w_1^2). \tag{22}$$

Hieraus und aus (9) und (10) folgt

$$\sum_{s=0}^{\infty} A_{2s}(-t_1^2, -w_1^2)\,(-1)^s \lambda^{2s} = \sum_{s=0}^{\infty} B_{2s}(t_1^2, w_1^2)\, \lambda^{2s}, \tag{23}$$

d. h. aber

$$(-1)^s A_{2s}(-t_1^2, -w_1^2) = B_{2s}(t_1^2, w_1^2). \tag{24}$$

Gehen wir nun in (14) von P zu P_1 über, so können wir dann auf Grund von (19) schreiben

$$S(-t_1^2, -w_1^2) = \sum_{s=0}^{\infty} A_{2s}(-t_1^2, -w_1^2)\, \sigma_s(-t_1^2) = -\sum_{s=0}^{\infty} (-1)^s A_{2s}(-t_1^2, -w_1^2)\, \tau_s(t_1^2). \tag{25}$$

Vertauschen wir jetzt in (7) w und t miteinander, so ersehen wir aus (24) und (25), daß

$$S(-t_1^2, -w_1^2) = -T(t_1^2, w_1^2) \tag{26}$$

sein wird.

Hieraus und aus (20) bis (22) folgt dann

$$\bar{M} = -N' + i\int_0^1 \sqrt{1-\lambda^2}\, f(-\lambda^2, -t_1^2, -w_1^2)\, d\lambda. \tag{27}$$

Wir gehen nun in (22) und (26) von P zu P_1 über und erhalten

$$f(-\lambda^2, w_2^2, t_2^2) = \psi(\lambda^2, -w_2^2 - t_2^2) \tag{28}$$

und

$$S(w_2^2, t_2^2) = -T(-w_2^2, -t_2^2). \tag{29}$$

Dasselbe tun wir in (21), woraus wegen (28) und (29) folgt

$$\bar{N}' = -w_2 w_1^3 S(w_2^2, t_2^2) + \frac{2}{\pi}\left\{\frac{i\pi}{2} - w_1 w_2 - \lg(w_1 + w_2)\right\}\int_0^1 \sqrt{1-\lambda^2}\, f(-\lambda^2, w_2^2, t_2^2)\, d\lambda. \tag{30}$$

Hieraus und aus (13) ergibt sich nun

$$\bar{N}' = -M + i\int_0^1 \sqrt{1-\lambda^2}\, f(-\lambda^2, w_2^2, t_2^2)\, d\lambda \tag{31}$$

und demnach wegen (27)

$$M - N' + i\int_0^1 \sqrt{1-\lambda^2}\, f(-\lambda^2, -t_1^2, -w_1^2)\, d\lambda = \bar{M} - \bar{N}' + i\int_0^1 \sqrt{1-\lambda^2}\, f(-\lambda^2, w_2^2, t_2^2)\, d\lambda, \tag{32}$$

und wir ersehen aus dem Vorhergehenden, daß jede Seite hiervon eine in bezug auf die X-Achse symmetrische Funktion darstellt.

Die Bedingung (22) führt demnach zu der Symmetrie in bezug auf die X-Achse, worüber wir in § 6 gesprochen haben. Führen wir außerdem noch die Bedingung (1) § 6 ein, so heben sich wegen (22) die beiden Integrale in (32) gegenseitig auf, und laut § 6 wird die Funktion $M - N'$ außerdem noch doppelpolig sein. Wir kommen später ausführlich darauf zurück.

Vertauschen wir nun in (11) w und t miteinander, d. h. gehen zu M' über, so erhalten wir in vollkommen analoger Weise unter derselben Bedingung (22) wegen (16)

$$M' - N - i\int_0^1 \sqrt{1-\lambda^2}\, f(-\lambda^2, -w_1^2, -t_1^2)\, d\lambda = \bar{M}' - N - i\int_0^1 \sqrt{1-\lambda^2}\, f(-\lambda^2, t_2^2, w_2^2)\, d\lambda, \tag{33}$$

wobei auch jetzt jede Seite von (33) eine in bezug auf die X-Achse symmetrische Funktion sein wird.

Wir können nicht direkt zu (33) aus (32) einfach durch Vertauschung von w und t miteinander kommen, wie dies aus dem Minuszeichen vor i in (33) erhellt. Dieses erklärt sich dadurch, daß bei dem Übergang von P zu P_1 bei M' wir unter dem lg, $-i$ haben werden, während bei demselben Übergang z. B. in M, unter dem lg, einfach i stehen wird.

Auf Grund ganz analoger Überlegungen erhalten wir aus (26) bis (29) § 5

$$(34) \qquad R = w_2 w_1^3 E(w_2^2, t_2^2) - w_1 w_2 \chi(w_2^2, w_2^2, t_2^2) + \frac{2}{\pi}\{w_1 w_2 + \lg(w_1 + w_2)\}\int_0^1 \frac{\chi(-\lambda^2, w_2^2, t_2^2)\, d\lambda}{\sqrt{1-\lambda^2}}$$

und

$$(35) \qquad W = w_1 w_2^3 C(w_1^2, t_1^2) + w_1 w_2 \gamma(w_1^2, w_1^2, t_1^2) - \frac{2}{\pi}\{w_1 w_2 - \lg(w_1 + w_2)\}\int_0^1 \frac{\gamma(-\lambda^2, w_1^2, t_1^2)\, d\lambda}{\sqrt{1-\lambda^2}},$$

wobei $E(w_2^2, t_2^2)$ und $C(w_1^2, t_1^2)$, ähnlich wie (14) und (17), ganze Funktionen von w_2^2 und t_2^2 resp. von w_1^2 und t_1^2 bedeuten.

Ferner ergeben sich, bei Einführung der zu (22) analogen Bedingung

$$(36) \qquad \chi(-\lambda^2, -t_1^2, -w_1^2) = \gamma(\lambda^2, t_1^2, w_1^2),$$

die Beziehungen

$$(37) \qquad R + W' + i\int_0^1 \frac{\chi(-\lambda^2, -t_1^2, -w_1^2)\, d\lambda}{\sqrt{1-\lambda^2}} = \overline{R} + \overline{W}' + i\int_0^1 \frac{\chi(-\lambda^2, w_2^2, t_2^2)\, d\lambda}{\sqrt{1-\lambda^2}}$$

und

$$(38) \qquad R' + W - i\int_0^1 \frac{\chi(-\lambda^2, -t_1^2, -w_1^2)\, d\lambda}{\sqrt{1-\lambda^2}} = \overline{R}' + \overline{W} - i\int_0^1 \frac{\chi(-\lambda^2, w_2^2, t_2^2)\, d\lambda}{\sqrt{1-\lambda^2}},$$

welche den Ausdrücken (32) und (33) entsprechen und wobei ebenfalls beide Seiten, in bezug auf die X-Achse, symmetrische Funktionen sein werden.

Bezüglich des Minuszeichens vor i in (38) gilt dieselbe Bemerkung wie bei (32).

Wir möchten noch darauf hinweisen, daß wir hier überall die Regel am Schluß der Bemerkung I § 3 angewandt haben.

§ 8.

Einige Folgerungen und Einführung von doppelpoligen Integralen der Wellengleichung.

Aus (1) und (3) § 5 ergibt sich, wegen (22) § 3, daß wir auf dem Streifen haben werden

$$(1) \qquad M + M' = 0; \qquad \xi = x = 0$$

Dieses Resultat folgt auch aus den entsprechenden Reihen für M und M' unter Beachtung des Umstandes, daß wegen (3) § 2

$$\lg(t_1 - t_2) = -\lg(t_1 + t_2) \tag{2}$$

sein wird.

Ganz analog folgt aus (26) und (28) § 5

$$R + R' = 0; \quad \xi = x = 0. \tag{3}$$

Nun könnte man auf Grund von (22) § 3 auf den ersten Blick schließen, daß die Summe der Integrale (2) und (4) § 5 auf dem Streifen nicht verschwinden wird. Wenden wir uns aber zu den Reihen (16) und (21) § 7, so ersehen wir, wieder wegen (22) § 3, daß

$$N + N' = 0; \quad \xi = x = 0 \tag{4}$$

ist. Auf analoge Weise folgt

$$W + W' = 0; \quad \xi = x = 0. \tag{5}$$

Funktionentheoretisch läßt sich dies folgendermaßen erklären. Wir denken uns in der Fig. 4 w_1 statt w_2 abgebildet und legen dann diese auf die Fig. 5 auf. Die Strecke 0 bis $+1$ entspricht dann der Strecke $A_1 A_2$ der Fig. 3. Die Integration bei N geht von 1 nach oben und bei N von $+1$ nach unten. Bei der Annäherung an den Streifen gelangen wir deshalb an die Strecke 0, $+1$ von verschiedenen Seiten. Beachten wir nun, daß $+1$ ein Verzweigungspunkt der Wurzel ist und fassen in diesem Fall die Strecke -1, $+1$ als Verzweigungsschnitt auf, so wird das Obige klar und stimmt auch mit den verschiedenen Vorzeichen in (16) und (25) § 5 überein.

Nach diesen allgemeinen Folgerungen wollen wir nun die Funktionen f und χ spezialisieren und setzen

$$f(\lambda^2, w_2^2, t_2^2) = \varepsilon(\lambda^2) J_0(2q\sqrt{t_1^2 + \lambda^2}\sqrt{w_2^2 - \lambda^2}); \tag{5}$$

wir erhalten demnach aus (1) § 5

$$M = \int_0^{w_2} \sqrt{\lambda^2 + 1}\, \varepsilon(\lambda^2) J_0(2q\sqrt{t_1^2 + \lambda^2}\sqrt{w_2^2 - \lambda^2})\, d\lambda. \tag{6}$$

Hieraus ergibt sich auf Grund der Bemerkung am Schluß des vorigen Paragraphen und unter Anwendung von (3) § 2

$$N = \int_1^{w_1} \sqrt{\lambda^2 - 1}\, \varepsilon(-\lambda^2) J_0(2q\sqrt{t_2^2 + \lambda^2}\sqrt{w_1^2 - \lambda^2})\, d\lambda. \tag{7}$$

Demnach also

$$M' = \int_0^{t_2} \sqrt{\lambda^2 + 1}\, \varepsilon(\lambda^2) J_0(2q\sqrt{w_1^2 + \lambda^2}\sqrt{t_2^2 - \lambda^2})\, d\lambda \tag{8}$$

und

$$N' = \int_1^{t_1} \sqrt{\lambda^2 - 1}\, \varepsilon(-\lambda^2) J_0(2q\sqrt{w_2^2 + \lambda^2}\sqrt{t_1^2 - \lambda^2})\, d\lambda. \tag{9}$$

Aus (6) erhalten wir dann

$$(10)\quad f(-\lambda^2, -t_1^2, -w_1^2) = f(-\lambda^2, w_2^2, t_2^2) = \varepsilon(-\lambda^2)\, J_0(2q\sqrt{t_1^2-\lambda^2}\sqrt{w_2^2+\lambda^2}).$$

Vergleichen wir dies mit (22) § 7, so ersehen wir, daß nun die Bedingung (1) § 6 erfüllt sein wird, d. h. die Bedingung für die Doppelpoligkeit.

Es heben sich daraufhin die Integrale auf beiden Seiten von (32) und (33) § 7 gegenseitig auf, und wir erhalten

$$(11)\quad u_1 = M - N',$$

$$(12)\quad u_2 = M' - N,$$

wobei also diese Funktionen, entsprechend den Paragraphen 6 und 7 und unter Berücksichtigung von § 4, doppelpolige und in bezug auf die X-Achse symmetrische *Integrale der Wellengleichung* sein werden.

Aus den entsprechenden Reihen erhalten wir

$$(13)\quad \begin{aligned} u_1 = {} & w_2 w_1^3 S(w_2^2, t_2^2) - t_1 t_2^3 T(t_1^2, w_1^2) \\ & + \frac{2}{\pi}\{w_1 w_2 - t_1 t_2 + \lg(w_1+w_2)(t_1+t_2)\} \\ & \times \int_0^1 \sqrt{1-\lambda^2}\, \varepsilon(-\lambda^2)\, J_0(2q\sqrt{t_1^2-\lambda^2}\sqrt{w_2^2+\lambda^2})\, d\lambda \end{aligned}$$

und

$$(14)\quad \begin{aligned} u_2 = {} & t_2 t_1^3 S(t_2^2, w_2^2) - w_1 w_2^3 T(w_1^2, t_1^2) \\ & + \frac{2}{\pi}\{t_1 t_2 - w_1 w_2 + \lg(w_1+w_2)(t_1+t_2)\} \\ & \times \int_0^1 \sqrt{1-\lambda^2}\, \varepsilon(-\lambda^2)\, J_0(2q\sqrt{w_1^2-\lambda^2}\sqrt{t_2^2+\lambda^2})\, d\lambda. \end{aligned}$$

Hierbei können wir von (13) zu (14) einfacher durch Vertauschung von w und t miteinander übergehen.

Aus (13) und (14) ersehen wir tatsächlich klar die angeführten Eigenschaften der Funktionen u_1 und u_2 unter Beachtung der Bemerkung bei (3) § 6.

Ferner setzen wir

$$(15)\quad \chi(\lambda^2, w_2^2, t_2^2) = \nu(\lambda^2)\, J_0(2q\sqrt{t_1^2+\lambda^2}\sqrt{w_2^2-\lambda^2}).$$

Demnach folgt aus (26) § 5

$$(16)\quad R = \int_0^{w_2} \frac{\nu(\lambda^2)\, J_0(2q\sqrt{t_1^2+\lambda^2}\sqrt{w_2^2-\lambda^2})\, d\lambda}{\sqrt{\lambda^2+1}}$$

und hieraus und aus (27) § 5, unter Beachtung der Bemerkung am Schluß von § 7 und wegen (3) § 2,

$$(17)\quad W = \int_1^{w_1} \frac{\nu(-\lambda^2)\, J_0(2q\sqrt{t_2^2+\lambda^2}\sqrt{w_2^2-\lambda^2})\, d\lambda}{\sqrt{\lambda^2-1}}.$$

Folglich ist

$$R' = \int_0^{t_2} \frac{\nu(\lambda^2)\, J_0\left(2q\sqrt{w_1^2+\lambda^2}\,\sqrt{t_2^2-\lambda^2}\right) d\lambda}{\sqrt{\lambda^2+1}} \tag{18}$$

und

$$W' = \int_1^{t_1} \frac{\nu(-\lambda^2)\, J_0\left(2q\sqrt{w_2^2+\lambda^2}\,\sqrt{t_1^2-\lambda^2}\right) d\lambda}{\sqrt{\lambda^2-1}}. \tag{19}$$

Aus (15) fließt

$$\chi(-\lambda^2, -t_1^2, -w_1^2) = \chi(-\lambda^2, w_2^2, t_2^2) = \nu(-\lambda^2)\, J_0(2q\sqrt{t_1^2-\lambda^2}\,\sqrt{w_2^2+\lambda} \tag{20}$$

d. h. es ist wegen (36) § 7 die Bedingung (2) § 6 der Doppelpoligkeit erfüllt und es sind demnach

$$u_3 = R + W' \tag{21}$$

und

$$u_4 = R' + W \tag{22}$$

Funktionen von derselben Art wie oben u_1 und u_2.

Aus den entsprechenden Reihen erhalten wir endlich

$$\begin{aligned} u_3 = {} & w_2 w_1^3 E(w_2^2, t_2^2) - w_1 w_2 \nu(w_2^2) + t_1 t_2^3 C(t_1^2, w_1^2) + t_1 t_2 \nu(-t_1^2) \\ & + \frac{2}{\pi}\{w_1 w_2 - t_1 t_2 + \lg(w_1+w_2)(t_1+t_2)\} \\ & \times \int_0^1 \frac{\nu(-\lambda^2)\, J_0\left(2q\sqrt{w_2^2+\lambda^2}\,\sqrt{t_1^2-\lambda^2}\right) d\lambda}{\sqrt{1-\lambda^2}} \end{aligned} \tag{23}$$

und

$$\begin{aligned} u_4 = {} & t_2 t_1^3 E(t_2^2, w_2^2) - t_1 t_2 \nu(t_2^2) + w_1 w_2^3 C(w_1^2, t_1^2) + w_1 w_2 \nu(-w_1^2) \\ & + \frac{2}{\pi}\{t_1 t_2 - w_1 w_2 + \lg(w_1+w_2)(t_1+t_2)\} \\ & \times \int_0^1 \frac{\nu(-\lambda^2)\, J_0\left(2q\sqrt{w_1^2-\lambda^2}\,\sqrt{t_2^2+\lambda^2}\right) d\lambda}{\sqrt{1-\lambda^2}}. \end{aligned} \tag{24}$$

Es ist leicht zu sehen, daß alle Integrale u_1, u_2, u_3 und u_4 in den beiden Polen A_1 und A_2 endlich sein werden, wie in § 6 gefordert.

§ 9.

Weitere Folgerungen und Zusammenstellung.

Auf Grund von (5) § 1 und der Paragraphen 2 und 3 erhalten wir

$$\frac{\partial}{\partial \varkappa x} = \frac{i}{4qw_1}\frac{\partial}{\partial w_1} - \frac{i}{4qt_1}\frac{\partial}{\partial t_1}. \tag{1}$$

Hieraus und unter Berücksichtigung von (5) und (6) § 2 ergibt sich leicht

$$\frac{\partial J_0\left(2q\sqrt{t_1^2+\lambda^2}\sqrt{w_2^2-\lambda^2}\right)}{\partial x\, x} = -\frac{i}{4q}\frac{\partial J_0\{\ \}}{\lambda\,\partial\lambda} \tag{2}$$

und

$$\frac{\partial J_0\left(2q\sqrt{w_2^2+\lambda^2}\sqrt{t_1^2-\lambda^2}\right)}{\partial x\, x} = \frac{i}{4q}\frac{\partial J_0\{\ \}}{\lambda\,\partial\lambda}\,. \tag{3}$$

Demnach folgt aus (6), (9) und (11) § 8

$$\frac{\partial u_1}{\partial x\, x} = \frac{i\,w_1}{4\,q\,w_2}\,\varepsilon(w_2^2) - \frac{i}{4q}\int_0^{w_2}\sqrt{\lambda^2+1}\,\varepsilon(\lambda^2)\frac{\partial J_0\left(2q\sqrt{t_1^2+\lambda^2}\sqrt{w_2^2-\lambda^2}\right)d\lambda}{\lambda\,\partial\lambda} \tag{4}$$

$$-\frac{i\,t_2}{4\,q\,t_1}\,\varepsilon(-t_1^2) - \frac{i}{4q}\int_1^{t_1}\sqrt{\lambda^2-1}\,\varepsilon(-\lambda^2)\frac{\partial J_0\left(2q\sqrt{w_2^2+\lambda^2}\sqrt{t_1^2-\lambda^2}\right)d\lambda}{\lambda\,\partial\lambda}$$

und in analoger Weise

$$\frac{\partial u_2}{\partial x\, x} = -\frac{i\,t_1}{4\,q\,t_2}\,\varepsilon(t_2^2) + \frac{i}{4q}\int_0^{t_2}\sqrt{\lambda^2+1}\,\varepsilon(\lambda^2)\frac{\partial J_0\left(2q\sqrt{w_1^2+\lambda^2}\sqrt{t_2^2-\lambda^2}\right)d\lambda}{\lambda\,\partial\lambda} \tag{5}$$

$$+\frac{i\,w_2}{4\,q\,w_1}\,\varepsilon(-w_1^2) + \frac{i}{4q}\int_1^{w_1}\sqrt{\lambda^2-1}\,\varepsilon(-\lambda^2)\frac{\partial J_0\left(2q\sqrt{t_2^2+\lambda^2}\sqrt{w_1^2-\lambda^2}\right)d\lambda}{\lambda\,\partial\lambda}$$

Auf Grund der allgemeinen Bemerkung am Anfang von § 8 und unter Beachtung von (16) § 3 erhalten wir die folgenden Resultate:

$$\left.\begin{aligned} u_1 + u_2 &= 0 \quad (6)\\ \frac{\partial(u_1+u_2)}{\partial x\, x} &\neq 0 \quad (7)\end{aligned}\right\}\ x = \xi = 0$$

und

$$\left.\begin{aligned} u_1 - u_2 &\neq 0 \quad (8)\\ \frac{\partial(u_1-u_2)}{\partial x\, x} &= 0 \quad (9)\end{aligned}\right\}\ x = \xi = 0.$$

Analog ergibt sich

$$\left.\begin{aligned} u_3 + u_4 &= 0 \quad (10)\\ \frac{\partial(u_3+u_4)}{\partial x\, x} &\neq 0 \quad (11)\end{aligned}\right\}\ x = \xi = 0$$

und

$$\left.\begin{aligned} u_3 - u_4 &\neq 0 \quad (12)\\ \frac{\partial(u_3-u_4)}{\partial x\, x} &= 0 \quad (13)\end{aligned}\right\}\ x = \xi = 0.$$

Wir untersuchen nun die oben gewonnenen Integrale der Wellengleichung in bezug auf die Symmetrie gegenüber der Y-Achse.

Da wir hierbei, wegen (15) § 3, z. B. von w_2 zu $-t_2$ resp. von w_1 zu $-t_1$ übergehen müssen, so werden wir uns bei dieser Betrachtung an die

Reihen halten, um Zweideutigkeiten zu vermeiden, da wir dabei gewissermaßen von der Fig. 4 zu der Fig. 5 überspringen.

Wir erhalten nun die folgenden Resultate, die sich leicht aus den entsprechenden Reihen ablesen lassen:

(14) $u_1 + u_2 =$ symmetrisch in bezug auf die Y-Achse.

Infolgedessen ergibt sich direkt auf Grund der allgemein bekannten Eigenschaften der Wellengleichung

$$(15) \qquad \frac{\partial(u_1+u_2)}{\partial x\, x} = 0; \quad x = 0; \quad \delta = \pm\frac{\pi}{2},$$

d. h. auf der Y-Achse, außerhalb des Streifens, also dort, wo $u_1 + u_2$ stetig sein wird.

Weiter folgt

(16) $u_1 - u_2 =$ antisymmetrisch in bezug auf die Y-Achse

und

$$\left.\begin{aligned} (17) \qquad u_1 - u_2 &= 0 \\ (18) \qquad \frac{\partial(u_1-u_2)}{\partial x\, x} &+ 0 \end{aligned}\right\} \quad x = 0; \quad \delta = \pm\frac{\pi}{2}.$$

Ganz analog erhalten wir

(19) $u_3 + u_4 =$ symmetrisch in bezug auf die Y-Achse

und demnach

$$(20) \qquad \frac{\partial(u_3+u_4)}{\partial x\, x} = 0; \quad x = 0; \quad \delta = \pm\frac{\pi}{2}.$$

Endlich folgt

(21) $u_3 - u_4 =$ antisymmetrisch in bezug auf die Y-Achse

und

$$\left.\begin{aligned} (22) \qquad u_3 - u_4 &= 0 \\ (23) \qquad \frac{\partial(u_3-u_4)}{\partial x\, x} &+ 0 \end{aligned}\right\} \quad x = 0; \quad \delta = \pm\frac{\pi}{2}.$$

Die Beziehungen (6) bis (13) einerseits und die Beziehungen (14) bis (23) anderseits entsprechen einander vollkommen.

§ 10.

Übergang zu $a = 0$.

Wir betrachten jetzt z. B. das Integral M in (6) § 8 und setzen dort $\frac{\lambda}{\sqrt{q}}$ statt λ, weshalb

$$(1) \qquad M = \int_0^{\sqrt{q}\,w_2} \sqrt{\lambda^2 + q}\,\mu(\lambda^2, q)\, J_0(2\sqrt{q t_1^2 + \lambda^2}\,\sqrt{q w_2^2 - \lambda^2})\, d\lambda$$

folgt, wobei

$$\mu(\lambda^2, q) = \frac{1}{q}\varepsilon\left(\frac{\lambda^2}{q}\right) \tag{2}$$

bedeutet. Da nun μ eine ganze Funktion von λ^2 ist, so ergibt sich, daß, falls μ bei $a = 0$ endlich bleibt, hierbei M unter Beachtung der Bemerkung II § 3 übergeht in

$$M_{a=0} = \int_0^{\sqrt{q}w_1} \lambda\,\mu(\lambda^2, 0)\, J_0(2\sqrt{q t_1^2 + \lambda^2}\,\sqrt{q w_1^2 - \lambda^2})\, d\lambda, \tag{3}$$

wo wir wieder mit q und w_1 rechnen können. Hierbei verliert M seine Zweiwertigkeit, entsprechend der Schleife III Fig. 3, da beide Pole zusammenfallen.

Vollkommen analoge Schlüsse beziehen sich auch auf die Integrale N, M' und N' in (7) bis (9) § 8.

Anders verhält es sich mit den Integralen (16) bis (19) § 8. Wir betrachten z. B. R in (16) § 8 und erhalten, analog (1):

$$R = \int_0^{\sqrt{q}w_1} \nu\left(\frac{\lambda^2}{q}\right) \frac{J_0(2\sqrt{q t_1^2 + \lambda^2}\,\sqrt{q w_2^2 - \lambda^2})\, d\lambda}{\sqrt{\lambda^2 + q}}. \tag{4}$$

Wegen der Eigenschaft von ν können wir schreiben

$$\nu\left(\frac{\lambda^2}{q}\right) = \sum_{s=0}^{\infty} c_{2s} \cdot \frac{\lambda^2}{q}, \tag{5}$$

wobei die c_{2s} im allgemeinen Funktionen von q sein werden.

Bleibt nun (5) endlich bei $q = 0$, so gilt alles bei M Gesagte für alle Glieder von R, für welche $s \geqq 1$ ist. Das Glied mit c_0 wird für $a = 0$ unendlich wie $\lg a$.

Wir untersuchen deshalb als Beispiel die Integrale

$$P_3 = R + W' = \int_0^{w_2} \frac{J_0(2q\sqrt{t_1^2 + \lambda^2}\,\sqrt{w_2^2 - \lambda^2})\, d\lambda}{\sqrt{\lambda^2 + 1}} + \int_1^{t_2} \frac{J_0(2q\sqrt{w_2^2 + \lambda^2}\,\sqrt{t_1^2 - \lambda^2})\, d\lambda}{\sqrt{\lambda^2 - 1}} \tag{6}$$

und

$$P_4 = R' + W = \int_0^{t_2} \frac{J_0(2q\sqrt{w_1^2 + \lambda^2}\,\sqrt{t_2^2 - \lambda^2})\, d\lambda}{\sqrt{\lambda^2 + 1}} + \int_1^{w_1} \frac{J_0(2q\sqrt{t_2^2 + \lambda^2}\,\sqrt{w_1^2 - \lambda^2})\, d\lambda}{\sqrt{\lambda^2 - 1}}. \tag{7}$$

Diese Integrale entsprechen vollkommen den Integralen u_3 und u_4 § 8, falls wir dort $\nu(\lambda^2) = 1$ setzen, und haben dieselben Eigenschaften, solange $a \geqq 0$ sein wird.

Indem wir die Integrationsvariable ändern ($\lambda^2 + 1 = y^2$ setzen und hinterher λ statt y schreiben), erhalten wir aus dem ersten Integral in (6)

das zweite Integral in (7) und aus dem zweiten Integral in (7) das erste Integral in (6). Demnach also statt (6)

$$(8)\quad P_3 = W + W' = \int_1^{w_1} \frac{J_0\left(2q\sqrt{t_2^2+\lambda^2}\sqrt{w_1^2-\lambda^2}\right)d\lambda}{\sqrt{\lambda^2-1}} + \int_1^{t_1} \frac{J_0\left(2q\sqrt{w_2^2+\lambda^2}\sqrt{t_1^2-\lambda^2}\right)d\lambda}{\sqrt{\lambda^2-1}}$$

und

$$(9)\qquad P_3 = P_4 .$$

Aus der Reihe (35) § 7 und der entsprechenden Reihe für W' und unter Beachtung, daß

$$(10)\qquad \int_0^1 \frac{J_0\left(2q\sqrt{t_2^2+\lambda^2}\sqrt{w_1^2-\lambda^2}\right)d\lambda}{\sqrt{1-\lambda^2}} = \int_0^1 \frac{J_0\left(2q\sqrt{w_2^2+\lambda^2}\sqrt{t_1^2-\lambda^2}\right)d\lambda}{\sqrt{1-\lambda^2}}$$

sein wird, erhalten wir

$$(11)\qquad P_3 = w_1 w_2^3 C(w_1^2, t_1^2) + t_1 t_1^3 C(t_1^2, w_1^2) + w_1 w_2 + t_1 t_2$$
$$- \frac{2}{\pi}\{w_1 w_2 + t_1 t_2 - \lg(w_1+w_2)(t_1+t_2)\}\int_0^1 \frac{J_0\left(2q\sqrt{t_2^2+\lambda^2}\sqrt{w_1^2-\lambda^2}\right)d\lambda}{\sqrt{1-\lambda^2}}.$$

Hieraus ersehen wir, daß P_3 auf dem Streifen verschwindet, wie es auch laut (6) § 9 und (9) sein muß.

Berechnet man nun tatsächlich (11) (wir unterdrücken die ziemlich langwierigen Rechnungen), so ergibt sich, falls wir laut der Bemerkung II § 3 überall $\frac{\sqrt{q}\,w_1}{\sqrt{q}}$ statt w_1 usw. schreiben, daß alle Glieder, mit Ausnahme desjenigen mit lg, bei $a = 0$ *endlich* bleiben.

Nun können wir schreiben

$$(12)\qquad \lg(w_1+w_2)(t_1+t_2) = \lg p(\sqrt{q}\,w_1 + \sqrt{q}\,w_2)(\sqrt{q}\,t_1 + \sqrt{q}\,t_2) - \lg pq,$$

wobei wir unter p eine beliebige von Null verschiedene Zahl verstehen.

Laut der Bemerkung II § 3 bleibt das erste Glied rechterseits in (12) bei $a = 0$ endlich. Deshalb folgt aus (12)

$$(13)\qquad \left\{\frac{-\log(w_1+w_2)(t_1+t_2)}{\lg pq}\right\}_{a=0} = 1.$$

Beachten wir nun ferner, daß

$$(14)\qquad \left\{\int_0^1 \frac{J_0\left(2q\sqrt{w_2^2+\lambda^2}\sqrt{t_1^2-\lambda^2}\right)d\lambda}{\sqrt{1-\lambda^2}}\right\}_{a=0} = \frac{\pi}{2} J_0(2qw_1t_1)$$

sein wird, so erhalten wir auf Grund des Vorhergehenden

$$(15)\qquad \left\{\frac{-P_3}{\lg pq}\right\}_{a=0} = J_0(2qw_1t_1) = J_0(\varkappa r_1).$$

§ 11.

Beziehungen zwischen verschiedenen Integralen der Wellengleichung und Folgerungen.

Wir geben hier einige Beziehungen zwischen verschiedenen Integralen der Wellengleichung an, und zwar nur die Resultate, da die Ableitung ziemlich umständliche Rechnungen erfordert und auf einigen allgemeinen Betrachtungen beruht, die wir hier nicht anführen können.

Es wird sein

$$e^{\pm i\varkappa x} = J_0(2qw_1t_1)$$

$$+\int_0^{w_1}\left\{\pm 2q\lambda J_0(2iq\lambda^2)+\frac{1}{2}\frac{\partial J_0(2iq\lambda^2)}{\partial\lambda}\right\}J_0(2q\sqrt{t_1^2+\lambda^2}\sqrt{w_1^2-\lambda^2})\,d\lambda$$

$$+\int_0^{t_1}\left\{\mp 2q\lambda J_0(2iq\lambda^2)+\frac{1}{2}\frac{\partial J_0(2iq\lambda^2)}{\partial\lambda}\right\}J_0(2q\sqrt{w_1^2+\lambda^2}\sqrt{t_1^2-\lambda^2})\,d\lambda. \tag{1}$$

Für das Sommerfeldsche Integral A in (12) und (13) § 4 haben wir

$$A = -\sqrt{iq}\int_0^{w_1} e^{-2q\lambda^2} J_0(2q\sqrt{t_1^2+\lambda^2}\sqrt{w_1^2-\lambda^2})\,d\lambda$$

$$+ i\sqrt{iq}\int_0^{t_1} e^{+2q\lambda^2} J_0(2q\sqrt{w_1^2+\lambda^2}\sqrt{t_1^2-\lambda^2})\,d\lambda. \tag{2}$$

Bezeichnen wir

$$B = e^{+i\varkappa x}\int_0^{\sqrt{2\varkappa r_1}\cos\frac{1}{2}\varphi_1} e^{-i\lambda^2}\,d\lambda = -e^{q(w_1^2-t_1^2)}\int_0^{i\sqrt{iq}(w_1+it_1)} e^{-i\lambda^2}\,d\lambda, \tag{3}$$

so haben wir analog (2)

$$B = -i\sqrt{iq}\int_0^{w_1} e^{+2q\lambda^2} J_0(2q\sqrt{t_1^2+\lambda^2}\sqrt{w_1^2-\lambda^2})\,d\lambda$$

$$+\sqrt{iq}\int_0^{t_1} e^{-2q\lambda^2} J_0(2q\sqrt{w_1^2+\lambda^2}\sqrt{t_1^2-\lambda^2})\,d\lambda. \tag{4}$$

Es sei hier nur nebenbei vermerkt, daß $e^{\pm i\varkappa x}$, A und B auch durch Integrale vom Typus (9) § 4 bei $\nu = 0$ und $c = 0$ dargestellt werden können und auch durch solche Typen (11) § 4, bei welchen statt des Pluszeichens das Minuszeichen vor λ^2 steht.

Aus dem Ausdruck für $\sin\varkappa x$, der sich aus (1) ergibt, bilden wir nun doppelpolige Integrale, und zwar

$$u_1 = 2q\int_0^{w_2}\sqrt{\lambda^2+1}\,J_0(2iq\lambda^2)\,J_0(2q\sqrt{t_1^2+\lambda^2}\sqrt{w_2^2-\lambda^2})\,d\lambda$$

$$-2q\int_1^{t_1}\sqrt{\lambda^2-1}\,J_0(2iq\lambda^2)\,J_0(2q\sqrt{w_2^2+\lambda^2}\sqrt{t_1^2-\lambda^2})\,d\lambda \tag{5}$$

und

$$(6)\qquad u_2 = 2q\int_0^{t_2}\sqrt{\lambda^2+1}\,J_0(2iq\lambda^2)\,J_0(2q\sqrt{w_1^2+\lambda^2}\,\sqrt{t_2^2-\lambda^2})\,d\lambda$$

$$-2q\int_1^{w_1}\sqrt{\lambda^2-1}\,J_0(2iq\lambda^2)\,J_0(2q\sqrt{t_2^2+\lambda^2}\,\sqrt{w_1^2-\lambda^2})\,d\lambda,$$

welche den Typen (11) und (12) des § 8 entsprechen.

Aus (5) und (6) folgt für $a=0$, laut den Ausführungen des vorigen Paragraphen,

$$(7)\qquad u_1 = i\sin\varkappa x; \qquad u_2 = -i\sin\varkappa x; \qquad a=0.$$

Ferner ergibt sich

$$(8)\qquad \frac{1}{2}\left\{\frac{u_2-u_1}{\varkappa}\right\}_{\varkappa=0} = -a\left\{\int_0^{w_2}\sqrt{\lambda^2+1}\,d\lambda - \int_1^{t_1}\sqrt{\lambda^2-1}\,d\lambda\right\}$$

$$+a\left\{\int_0^{t_2}\sqrt{\lambda^2+1}\,d\lambda - \int_1^{w_1}\sqrt{\lambda^2-1}\,d\lambda\right\}$$

oder wegen (7) und (8) § 7 und (11) § 3

$$(9)\qquad \frac{1}{2}\left\{\frac{u_2-u_1}{\varkappa}\right\}_{\varkappa=0} = a\{t_1t_2 - w_1w_2\} = -ia\cosh\xi\cdot\cos\delta.$$

Zum Schluß noch ein Beispiel zu dem Übergang zu $a=\infty$ im Sinne der Bemerkung III § 3. Wir haben

$$(10)\qquad \left\{\sqrt{iq}\,e^{q(w_2^2-t_2^2)}\int_0^{w_2}e^{-2q\lambda^2}J_0(2q\sqrt{t_2^2+\lambda^2}\,\sqrt{w_2^2-\lambda^2})\,d\lambda\right\}_{a=\infty} = -\frac{\sqrt{\pi}}{2}e^{-\frac{i\pi}{4}}$$

und

$$(11)\qquad \left\{\sqrt{iq}\,e^{q(t_1^2-w_2^2)}\int_0^{w_2}e^{+2q\lambda^2}J_0(2q\sqrt{t_2^2+\lambda^2}\,\sqrt{w_2^2-\lambda^2})\,d\lambda\right\}_{a=\infty} = 0,$$

was mit Hilfe der Sommerfeldschen Integrale abgeleitet worden ist.

§ 12.

Betrachtungen zur Beugung an einem blanken und schwarzen Streifen.

Als einfallende Welle nehmen wir eine ebene Welle an, welche sich in Richtung der positiven X-Achse fortpflanzt. Ist die elektrische Kraft E_0 dieser Welle parallel der Z-Achse (Fall I), so setzen wir (indem wir den Zeitfaktor unterdrücken)

$$(1)\qquad E_0 = e^{-i\varkappa x}.$$

Ist die magnetische Kraft der einfallenden Welle H_0 parallel der Z-Achse (Fall II), so schreiben wir

$$(2)\qquad H_0 = \frac{e^{-i\varkappa x}}{c},$$

wo c die Lichtgeschwindigkeit bedeutet, alles im absoluten elektromagnetischem Maßsystem ausgedrückt.

Wir führen nun die folgenden Funktionen ein:

$$F_1 = \frac{e^{\frac{i\pi}{4}}}{\sqrt{\pi}} B - \frac{e^{+i\varkappa x}}{2} \tag{3}$$

und

$$F_2 = \frac{e^{\frac{i\pi}{4}}}{\sqrt{\pi}} A - \frac{e^{-i\varkappa x}}{2}, \tag{4}$$

wobei A und B die Sommerfeldschen Integrale in (2) und (3) des vorigen Paragraphen darstellen.

Die elektrische Kraft der gebeugten Welle bezeichnen wir für den ersten Fall durch E_1 und sie wird also parallel der Z-Achse sein. Im zweiten Fall bezeichnen wir durch H_1 die magnetische Kraft der gebeugten Welle, welche auch parallel der Z-Achse ist. Bezeichnen wir ferner durch Y die elektrische Kraft für den letzteren Fall, parallel der Y-Achse, so wird sein

$$Y = ic \frac{\partial H_1}{\partial \varkappa x}. \tag{5}$$

Wir gehen nun im Sinne der Bemerkung III § 3 zu $a = \infty$ über, d. h. zu einer Halbebene mit der Kante im Pole A_1 der Fig. 1.

Es lautet dann bekanntlich die Sommerfeldsche Lösung für die blanke Halbebene für den ersten Fall

$$E_1 = F_2 + F_1 \tag{6}$$

und für den zweiten Fall

$$H_1 = \frac{F_2 - F_1}{c}. \tag{7}$$

Ferner führen wir noch die Funktionen

$$G_1 = \frac{V_1}{2} - \frac{e^{+i\varkappa x}}{2} \tag{8}$$

und

$$G_2 = \frac{V_2}{2} - \frac{e^{-i\varkappa x}}{2} \tag{9}$$

ein, wobei V_1 und V_2 gewisse doppelpolige Integrale bedeuten, die der Lösung der Beugung an einem Streifen genügen sollen.

Wir setzen dann für den ersten Fall

$$E_1 = G_2 + G_1 = \frac{V_1 + V_2}{2} - \cos \varkappa x \tag{10}$$

und für den zweiten

$$H_1 = G_2 - G_1 = \frac{V_2 - V_1}{2c} + \frac{i \sin \varkappa x}{c}. \tag{11}$$

Aus den Ausführungen der beiden vorhergehenden Paragraphen ersehen wir, daß wir immer solche doppelpolige Integrale V_1 und V_2 bilden können, daß für $a = 0$, $V_1 = e^{+i\varkappa x}$ und $V_2 = e^{-i\varkappa x}$ werden.

Dann verschwinden E_1 und H_1 für $a = 0$, wie es ja auch sein muß, da ja hierbei kein Hindernis für die einfallende Welle vorhanden sein wird.

Auf dem Streifen wird $E_0 = 1$ sein. Wegen (7) § 11 und (6) und (10) § 9 wird dabei $E_1 = -1$. Im zweiten Fall ergibt sich analog und wegen (11) und (5) $Y = -1$. Dieses alles entspricht der Beugung an einem blanken Streifen bei der angenommenen Art der einfallenden Welle.

Um nun die für die Beugung notwendigen Funktionen explizite anzugeben, müssen wir noch das Verhalten unserer doppelpoligen Integrale in unendlicher Entfernung vom Koordinatenanfang untersuchen.

Wir hoffen darauf in einiger Zeit zurückzukommen.

Sind nun V_1 und V_2 die notwendigen Lösungen, so muß z. B. $\frac{V_1}{2}$ beim Übergang zu $a = \infty$, im Sinne der Bemerkung III § 3, in $\frac{e^{\frac{i\pi}{4}}}{\sqrt{\pi}} B$ in (3) übergehen.

Ferner muß $\frac{V_2 - V_1}{2}$ in (11) der Beziehung (9) § 11 genügen, was, wie aus § 11 ersichtlich, möglich ist. Ist nämlich (9), § 11 erfüllt, so ergeben uns (5) und (11)

$$Y_{x=0} = \left\{ ic \frac{\partial H_1}{\partial x x} \right\}_{x=0} = \frac{\partial}{\partial x} \{ a \cosh \xi \cdot \cos \delta - x \} \tag{12}$$

oder, wegen (2) § 1,

$$Y_{x=0} = a \frac{\partial e^{-\xi} \cos \delta}{\partial x}. \tag{13}$$

Dieses entspricht dem statischen Fall, wobei sich unser Streifen in einem homogenen elektrischen Feld befindet, dessen elektrische Kraft (von der Stärke 1) parallel der Y-Achse sein wird.

Der Zusatzpotential V für diesen Fall lautet

$$V = a e^{-\xi} \sin \delta, \tag{14}$$

und hieraus folgt für die entsprechende elektrische Kraft $Y_{x=0}$

$$Y_{x=0} = -\frac{\partial V}{\partial y} = -a \frac{\partial e^{-\xi} \sin \delta}{\partial y}. \tag{15}$$

Auf Grund von (5) und (6) § 1 ist es nun leicht, die Identität der rechten Seiten von (15) und (13) nachzuweisen. Diese Identität ist ebenfalls eine notwendige Bedingung für die Lösung, bestimmt aber dieselbe nicht eindeutig, wie aus dem vorhergehenden Paragraphen ersichtlich.

Zum Schluß noch folgende Bemerkung:

Angenommen wieder, es seien V_1 und V_2 die richtigen Lösungen. Außerdem nehmen wir den Streifen als schwarz an. Für den ersten Fall setzen wir dann als Lösung an

$$E_1 = G_2 = \frac{V_2}{2} - \frac{e^{-ixx}}{2}. \tag{16}$$

Die gesamte elektrische Kraft E wird sein

(17) $$E = E_1 + E_0.$$

Gehen wir nun von dem Punkt P der Fig. 3 längs der Schleife II herum, so gelangen wir zu einem Punkt P' im nichtphysikalischen Raum, welcher unterhalb P liegen wird. Die gesamte elektrische Kraft E' in P' ergibt sich dann, unter Berücksichtigung der Zweiwertigkeit von V_2, zu

(18) $$E' = E_0 + E_1' = E_0 - \frac{V_2}{2} - \frac{e^{-i\varkappa x}}{2}.$$

Hieraus und aus (16) und (17) erhalten wir

(19) $$E + E' = 2E_0 - e^{-i\varkappa x} = E_0,$$

und dies stimmt vollkommen mit den Ausführungen in meiner Annalenarbeit überein [3]).

Wir hoffen. auf die Untersuchung der Funktionen V_1 und V_2 und auf die Lösungen der Wellengleichung für mehr als zwei Pole noch an anderer Stelle zurückzukommen.

Leningrad, Januar 1928.

Zusatz während des Druckes. Statt des Ausdruckes für $\sin \varkappa x$, der sich aus (1) § 11 ergibt, können wir auch schreiben

(1) $$i \sin \varkappa x = 2q \int_0^{w_2} \lambda J_0(2qi\lambda\sqrt{\lambda^2+1})\, J_0(2q\sqrt{t_1^2+\lambda^2}\,\sqrt{w_2^2-\lambda^2})\, d\lambda$$
$$- 2q \int_1^{t_1} \lambda J_0(2qi\lambda\sqrt{\lambda^2-1})\, J_0(2q\sqrt{w_2^2+\lambda^2}\,\sqrt{t_1^2-\lambda^2})\, d\lambda,$$

gültig in der ganzen Ebene der Fig. 1. Durch Differentiation nach x erhalten wir den entsprechenden Ausdruck für $\cos \varkappa x$.

Wir führen nun folgende doppelpolige Integrale ein

(2) $$F_1 = 2q \int_0^{w_2} \frac{\lambda^2 + \frac{1}{2}}{\sqrt{\lambda^2+1}} J_0(2iq\lambda\sqrt{\lambda^2+1})\, J_0(2q\sqrt{t_1^2+\lambda^2})^2 \sqrt{w_2^2-\lambda^2})\, d\lambda$$
$$- 2q \int_1^{t_1} \frac{\lambda^2 - \frac{1}{2}}{\sqrt{\lambda^2-1}} J_0(2iq\lambda\sqrt{\lambda^2-1})\, J_0(2q\sqrt{w_2^2+\lambda^2})^2 \sqrt{t_1^2-\lambda^2})\, d\lambda$$

und

(3) $$D_1 = -2qi \int_0^{w_2} \lambda J_1(2iq\lambda\sqrt{\lambda^2+1})\, J_0(2q\sqrt{t_1^2+\lambda^2}\sqrt{w_2^2-\lambda^2})\, d\lambda$$
$$- 2qi \int_1^{t_1} \lambda J_1(2iq\lambda\sqrt{\lambda^2-1})\, J_0(2q\sqrt{w_2^2+\lambda^2}\sqrt{t_1^2-\lambda^2})\, d\lambda.$$

[3]) W. v. Ignatowsky, loc. cit. § 12 und 15.

Jetzt gehen wir (wie in der Fig. 4) zu den Polarkoordinaten ϱ und ϑ über mit dem Pol im Koordinatenanfang der Fig. 1.

Es lassen sich dann folgende Beziehungen nachweisen

$$U = F_1 + D_1 - e^{ikx} = \frac{U_0(\varrho, \vartheta)}{\sqrt{\varrho}}; \quad \varrho \to \infty \tag{4}$$

und

$$W = F_1 - D_1 + e^{-ikx} = \frac{W_0(\varrho, \vartheta)}{\sqrt{\varrho}}; \quad \varrho \to \infty, \tag{5}$$

gültig für gewisse Wertbereiche von ϑ. Hierbei bleiben U_0 und W_0 für $\varrho = \infty$ endlich.

Die Methode, welche zur Ableitung von (4) und (5) diente, ist ganz allgemein und läßt sich auf ganze Klassen unserer doppelpoligen Integrale anwenden. Damit wäre die Frage des Verhaltens dieser Integrale für $\varrho = \infty$ gewissermaßen erledigt. Für $\varrho \to \infty$ ist die Untersuchung schwieriger, da der Bau der Funktionen U_0 und W_0 resp. ähnlicher Funktionen ziemlich kompliziert ist, was weitere Betrachtungen erfordert, um diese letzteren Funktionen in eine bequemere Form zu bringen.

Bezüglich der Größenordnung von ϱ, welche zum Bestehen von (4) und (5) notwendig ist, sei bemerkt, daß wir nicht nur $\frac{\varrho}{a}$ als groß angenommen, sondern außerdem die Bedingung von K. Schwarzschild[4]) eingeführt haben.

Leningrad, Juni 1928.

---

[4]) K. Schwarzschild, Beugung und Polarisation des Lichtes, Math. Annalen **55** (1902), S. 230, Formel 112.

(Eingegangen am 2. Juli 1928.)

Quelques considérations sur la théorie des intégrales singulières.

Par

M. Paul Romanovski à Moscou.

Dans le présent travail sont étudiées certaines propriétés des intégrales singulières. Le signe $\int$ qui se rencontre dans le texte désigne toujours des intégrales au sens de M. Lebesgue.

1. La notion de noyau. — Considérons une fonction $\varphi(x, n)$ déterminée dans l'intervalle (a, b) et dépendant d'un paramètre naturel n (nous supposons que pour chaque valeur de n elle est bornée et sommable). Si, n croissant indéfiniment, les intégrales $\int_a^l \varphi(x,n)\,dx$, où $a < l \leqq b$, tendent vers une limite déterminée G indépendante de l et, de plus, y tendent uniformément pour $a + \nu \leqq l \leqq b$ quelque petit que soit le nombre positif ν, alors $\varphi(x, n)$ est appelée *noyau* (avec singularité en a), et le nombre G s'appelle la *valeur intégrale* de ce noyau.

La condition nécéssaire et suffisante pour que $\varphi(x, n)$ soit un noyau consiste en l'existence d'une limite de $\int_a^b \varphi(x, n)\,dx$ et en ce que $\int_b \varphi(x, n)\,dx$ tende uniformément vers zéro pour $a + \nu \leqq l \leqq b$, quelque petit que soit $\nu > 0$.

Une combinaison linéaire de noyaux est un noyau. Sa valeur intégrale est la même combinaison linéaire des valeurs intégrales des noyaux donnés.

En changeant le noyau $\varphi(x, n)$ arbitrairement pour un nombre fini de valeurs de n ou en le changeant pour chaque valeur de n dans un ensemble de mesure nulle, nous obtiendrons de nouveau un noyau avec la même valeur intégrale.

Un noyau se nomme à signe constant (positif ou négatif) s'il est toujours $\geqq 0$ ou toujours $\leqq 0$. Dans le cas contraire il s'appelle à signe variable.

La fonction considérée $\varphi(x, n)$ ne peut être supposée déterminée pour chaque valeur de n que presque partout sur (a, b).

2. Remarque sur la propriété fondamentale de noyau à signe constant. — La propriété fondamentale du noyau à signe constant, ainsi que l'on sait [1]), consiste en ce que si $\varphi(x, n)$ est un tel noyau (à valeur intégrale G) et $f(x)$ est une fonction bornée sommable (dans l'intervalle (a, b)) et $f(a+0)$ existe, alors $\varphi(x, n) f(x)$ est un noyau à valeur intégrale $G f(a+0)$.

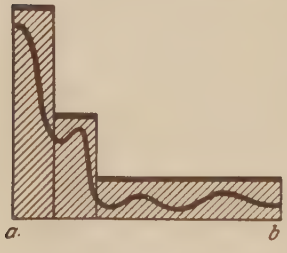
L'hypothèse que $f(x)$ soit bornée est essentielle. Sans cette hypothèse le théorème cesse d'être vrai. Montrons cela sur l'exemple suivant. Soit $a < \alpha_n < \beta_n < b$, où α_n et β_n tendent vers b de telle sorte que $\frac{b-\alpha_n}{b-\beta_n}$ croisse indéfiniment. Soit $\varphi(x, n)$ un noyau positif quelconque. Posons

$$\varphi_1(x, n) = \begin{cases} \varphi(x, n) & \text{en dehors de } (\alpha_n, \beta_n); \\ \dfrac{1}{\sqrt{b-x}} & \text{dans } (\alpha_n, \beta_n). \end{cases}$$

$f(x) = \frac{1}{\sqrt{b-x}}$ étant sommable dans (a, b), il est aisé de voir que $\varphi_1(x, n)$ est un noyau (positif). Or

$$\int_a^b f(x)\,\varphi_1(x, n)\,dx \geqq \int_{\alpha_n}^{\beta_n} f(x)\,\varphi_1(x, n)\,dx = \int_{\alpha_n}^{\beta_n} \frac{1}{b-x}\,dx = \log\frac{b-\alpha_n}{b-\beta_n} \to +\infty.$$

3. Sur deux types particuliers de noyaux positifs. — Nous soumettrons à une étude spéciale les deux types particuliers suivants de noyaux positifs.



Le premier type de noyaux positifs que nous mettons en évidence sont les noyaux positifs $\varphi(x, n)$ qui représentent pour chaque valeur de n une fonction non croissante de la variable x.

Le second type que nous mettons en évidence est formé de noyaux positifs $\varphi(x, n)$ possédant les propriétés: la borne supérieure de φ dans $(a+\nu, b)$ tend vers zéro avec $\frac{1}{n}$, quelque petit que soit $\nu > 0$; il est possible de choisir les nombres positifs $\varepsilon_n \to 0$ de façon qu'en désignant

[1]) Voir par exemple E. Picard, Traité d'Analyse, vol. I, troisième édition (Paris 1922), p. 266.

par M_n la borne supérieure de $\varphi(x, n)$ dans (a, b) et par m_n la borne supérieure de $\varphi(x, n)$ dans $(a + \varepsilon_n, b)$, les nombres $M_n \varepsilon_n$ et m_n soient bornés dans leur ensemble[2]).

4. Propriété des noyaux positifs du premier type (d'après la terminologie n° 3). — Pour un pareil noyau $\varphi(x, n)$ avec valeur intégrale G a lieu le théorème suivant.

Si la fonction $f(x)$, sommable dans l'intervalle (a, b), est telle que $\frac{1}{\varepsilon}\int_a^{a+\varepsilon} [f(x) - f(a)]\,dx$ tend vers zéro avec ε, alors

$$\int_a^l f(x)\,\varphi(x, n)\,dx \to G f(a)$$

uniformément pour toutes les valeurs de l qui satisfont aux inégalités $a + \nu \leqq l \leqq b$, quelque petit que soit $\nu > 0$.

Pour le démontrer écrivons

$$\int_a^l f(x)\,\varphi(x, n)\,dx = f(a)\int_a^l \varphi(x, n)\,dx + \int_a^l [f(x) - f(a)]\,\varphi(x, n)\,dx$$

et représentons le second terme sous la forme $\int_a^{a+\varepsilon'} + \int_{a+\varepsilon'}^l (\varepsilon' < \nu)$. Considérons d'abord $\int_{a+\varepsilon'}^l$. Comme

$$\left| \int_{a+\varepsilon'}^l \right| \leqq \int_{a+\varepsilon'}^l |f(x) - f(a)|\,\varphi(x, n)\,dx \leqq \varphi(a + \varepsilon', n) \int_a^b |f(x) - f(a)|\,dx,$$

il suffit, pour prouver que $\int_{a+\varepsilon'}^l$ tend vers zéro uniformément, de montrer que $\varphi(a + \varepsilon', n) \to 0$. Soit ε'' un nombre positif inférieur à ε', alors

$$\int_{a+\varepsilon''}^{a+\varepsilon'} \varphi(x, n)\,dx \geqq (\varepsilon' - \varepsilon'')\,\varphi(a + \varepsilon', n).$$

Or le premier membre de cette inégalité tend vers zéro; par conséquent le second membre y tend a fortiori, d'où il suit que $\varphi(a + \varepsilon', n) \to 0$ (nous voyons ainsi que le noyau non croissant tend vers zéro pour toute valeur de $x > a$). Considérons à présent $\int_a^{a+\varepsilon'}$. Si l'on suppose que le noyau

[2]) On voit aisément que l'on peut définir le noyau de ce type comme un noyau positif qui peut être majoré par une courbe à trois marches (voir la figure) qui a une aire bornée, la hauteur de la seconde marche bornée, la hauteur de la troisième marche infiniment petite et les longueurs de la première et seconde marche infiniment petites.

$\varphi(x, n)$ est absolument continu pour chaque valeur de n (dans quel cas il possède une dérivée sommable et est son intégrale indéfinie), alors une intégration par parties donne[3])

$$\int_a^{a+\varepsilon'} = \int_a^{a+\varepsilon'} [f(x) - f(a)]\,dx \cdot \varphi(a+\varepsilon', n) + \int_a^{a+\varepsilon'} \left\{ \int_a^x [f(x) - f(a)]\,dx \right\} [-\varphi'(x, n)]\,dx.$$

Écrivons le premier terme du second membre de cette formule sous forme de produit

$$\varepsilon' \varphi(a+\varepsilon', n) \cdot \frac{1}{\varepsilon'} \int_a^{a+\varepsilon'} [f(x) - f(a)]\,dx.$$

Le second facteur de ce produit par hypothèse tend vers zéro avec ε' et le premier est borné pour toutes les valeurs de ε' et n, ce qui découle des inégalités

$$\int_a^b \varphi(x, n)\,dx \geqq \int_a^{a+\varepsilon'} \varphi(x, n)\,dx \geqq \varepsilon' \varphi(a+\varepsilon', n).$$

Ainsi le premier terme de la formule d'intégration par parties tend vers zéro uniformément pour toutes les valeurs de n quand $\varepsilon' \to 0$. Écrivons le second terme de cette formule sous la forme

$$\int_a^{a+\varepsilon'} \left\{ \frac{1}{x-a} \int_a^x [f(x) - f(a)]\,dx \right\} [-(x-a)\varphi'(x, n)]\,dx.$$

Sa valeur absolue, en considération de ce que $\varphi'(x, n) \leqq 0$, ne surpasse pas

$$\max_{a \leqq x \leqq a+\varepsilon'} \left| \frac{1}{x-a} \int_a^x [f(x) - f(a)]\,dx \right| \cdot \int_a^{a+\varepsilon'} -(x-a)\varphi'(x, n)\,dx.$$

Dans ce produit le premier facteur, ainsi qu'on le voit d'après la condition, tend vers zéro avec ε' et le second est borné pour toutes les valeurs de ε' et de n, car

$$\int_a^{a+\varepsilon'} -(x-a)\varphi'(x, n)\,dx = [-(x-a)\varphi(x, n)]_a^{a+\varepsilon'} + \int_a^{a+\varepsilon'} \varphi(x, n)\,dx$$

$$= -\varepsilon' \varphi(a+\varepsilon', n) + \int_a^{a+\varepsilon'} \varphi(x, n)\,dx \leqq \int_a^{a+\varepsilon'} \varphi(x, n)\,dx \leqq \int_a^b \varphi(x, n)\,dx.$$

Par conséquent

$$\int_a^{a+\varepsilon'} \left\{ \int_a^x [f(x) - f(a)]\,dx \right\} [-\varphi'(x, n)]\,dx$$

---

[3]) Un artifice analogue était employé par M. J. Priwaloff pour le cas de l'intégrale de Poisson (Recueil mathématique de la Société mathématique de Moscou **30**, p. 314, en russe).

tend vers zéro uniformément pour toutes les valeurs de n quand $\varepsilon' \to 0$. De cette façon nous avons démontré que $\int_a^{a+\varepsilon'}$ tend vers zéro uniformément pour toutes les valeurs de n quand $\varepsilon' \to 0$; par conséquent quand ε' est suffisamment petit, $\left| \int_a^{a+\varepsilon'} \right| < \varepsilon$ pour toutes les valeurs de n (ε étant un nombre positif donné à l'avance).

5. Supplément au n° précédent. — Nous avons supposé le noyau $\varphi(x, n)$ absolument continu pour chaque valeur de n. Soit cette restriction levée.

Observons d'abord que si $\varphi(x, n)$ est un noyau quelconque avec valeur intégrale G et $\varphi_1(x, n)$ (bornée et sommable pour chaque valeur de n) est telle que

$$\int_a^b |\varphi_1(x, n) - \varphi(x, n)|\, dx$$

tend vers zéro avec $\frac{1}{n}$, alors $\varphi_1(x, n)$ est aussi un noyau de valeur intégrale G. Cela est visible sur l'égalité

$$\int_a^l \varphi_1(x, n)\, dx = \int_a^l \varphi(x, n)\, dx + \int_a^l [\varphi_1(x, n) - \varphi(x, n)]\, dx,$$

en observant que

$$\left| \int_a^l [\varphi_1(x, n) - \varphi(x, n)]\, dx \right| \leqq \int_a^b |\varphi_1(x, n) - \varphi(x, n)|\, dx.$$

Observons ensuite, que si $\varphi(x)$ est une fonction non croissante dans (a, b), alors il existe dans cette intervalle une fonction absolument continue non croissante $\varphi_1(x)$ telle que

$$\int_a^b |\varphi_1(x) - \varphi(x)|\, dx < \varepsilon,$$

quelque petit que soit le nombre positif ε. Si $\varphi(x)$ est continue on peut prendre pour $\varphi_1(x)$ un polygone inscrit dans $\varphi(x)$ à cotés suffisamment petits. On voit aisément ensuite que la fonction cherchée $\varphi_1(x)$ existe si $\varphi(x)$ n'a qu'un nombre fini de discontinuités. Il est clair ensuite que la fonction cherchée $\varphi_1(x)$ existe si le nombre de discontinuités est infini mais la somme des valeurs absolues de tous les sauts de la fonction $\varphi(x)$ est suffisamment petite. Enfin, dans le cas général, $\varphi(x)$ est représentable sous forme de somme d'une fonction non croissante à nombre fini de discontinuités et d'une fonction non croissante avec une somme aussi petite

que l'on veut de valeurs absolues des sauts; par conséquent dans le cas général la fonction $\varphi_1(x)$ existe aussi. Il est évident que l'on peut toujours supposer $\varphi(a) \geqq \varphi_1(x) \geqq \varphi(b)$ pour toutes les valeurs de x dans (a, b).

En se basant sur ces deux remarques il est facile, pour tout noyau positif non croissant $\varphi(x, n)$, de construire un noyau non croissant positif absolument continu $\varphi_1(x, n)$ de telle sorte que de ce que le théorème n° 4 est valable pour $\varphi_1(x, n)$ découle la valadité de ce théorème pour $\varphi(x, n)$.

En effet, écrivons

$$\int_a^l f(x)\varphi(x,n)\,dx = \int_a^l f(x)\varphi_1(x,n)\,dx + \int_a^l f(x)[\varphi(x,n) - \varphi_1(x,n)]\,dx.$$

Si $f(x)$ est bornée ($\leqq M$ en valeur absolue), alors la valeur absolue du second terme est $\leqq M\int_a^b |\varphi - \varphi_1|\,dx$; par conséquent ce second terme tend vers zéro uniformément si $\varphi_1(x, n)$ est construit de façon que $\int_a^b |\varphi - \varphi_1|\,dx \to 0$. Si $f(x)$ est seulement sommable, mais le noyau $\varphi(x, n)$ est continu pour chaque valeur de n, alors le dit second terme tend vers zéro uniformément si l'on prend $\varphi_1(x, n)$ de façon que $\varphi - \varphi_1 \to 0$ uniformément, car la valeur absolue du dit terme est $\leqq \max|\varphi - \varphi_1| \cdot \int_a^b |f(x)|\,dx$.

Enfin, dans le cas général le choix de $\varphi_1(x, n)$ doit être effectué individuellement pour chaque $f(x)$. On peut, par exemple, le faire de la façon suivante. La fonction sommable $f(x)$ est toujours représentable sous forme d'une somme d'une fonction bornée sommable $f_1(x)$ et d'une fonction $f_2(x)$ telle que $\int_a^b |f_2|\,dx$ soit aussi petite qu'on le veut. En posant $f(x) = f_{1n}(x) + f_{2n}(x)$, où f_{1n} est bornée et $\int_a^b |f_{2n}|\,dx < \varepsilon_n$, désignons par M_n la borne supérieure de $|f_{1n}|$. Nous avons alors

$$\left|\int_a^l f(x)[\varphi(x,n) - \varphi_1(x,n)]\,dx\right| \leqq M_n \int_a^b |\varphi - \varphi_1|\,dx + \varepsilon_n \max|\varphi - \varphi_1|$$

$$\leqq M_n \int_a^b |\varphi - \varphi_1|\,dx + \varepsilon_n \varphi(a, n),$$

or cela sera $< \frac{2}{n}$ si l'on prend $\varepsilon_n = \frac{1}{n\varphi(a,n)}$ et si l'on dispose du noyau $\varphi_1(x, n)$ de telle sorte que $\int_a^b |\varphi - \varphi_1|\,dx < \frac{1}{nM_n}$.

6. Propriété des noyaux positifs du second type (d'après la terminologie n° 3). — Pour un tel noyau $\varphi(x, n)$ avec valeur intégrale G, le théorème suivant a lieu.

Si la fonction $f(x)$, sommable dans l'intervalle (a, b), est telle que $\frac{1}{\varepsilon}\int_a^{a+\varepsilon} |f(x) - f(a)|\,dx$ tend vers zéro avec ε, alors

$$\int_a^l f(x)\,\varphi(x, n)\,dx \to G f(a)$$

uniformément pour toutes les valeurs de l qui satisfont aux inégalités $a + \nu \leqq l \leqq b$, quelque petit que soit $\nu > 0$.

Commençant la démonstration comme dans n° 4, considérons d'abord $\int_a^{a+\varepsilon'}$. Conservant les notations n° 3 écrivons $\int_a^{a+\varepsilon'} = \int_a^{a+\varepsilon_n} + \int_{a+\varepsilon_n}^{a+\varepsilon'}$. Ici

$$\left|\int_a^{a+\varepsilon_n}\right| \leqq M_n \int_a^{a+\varepsilon_n} |f(x) - f(a)|\,dx = M_n \varepsilon_n \cdot \frac{1}{\varepsilon_n} \int_a^{a+\varepsilon_n} |f(x) - f(a)|\,dx,$$

par conséquent $\int_a^{a+\varepsilon_n} \to 0$. Ensuite

$$\left|\int_{a+\varepsilon_n}^{a+\varepsilon'}\right| \leqq m_n \int_a^{a+\varepsilon'} |f(x) - f(a)|\,dx,$$

donc $\int_{a+\varepsilon_n}^{a+\varepsilon'}$ est uniformément petit pour toutes les valeurs de n si ε' est suffisamment petit. Il suit de là que, quelque petit que soit le nombre positif ε,

$$\left|\int_a^{a+\varepsilon'}\right| < \varepsilon$$

pour ε' convenablement choisi et n suffisamment grand. Il reste encore à montrer que $\int_{a+\varepsilon'}^{l} \to 0$ uniformément pour les valeurs de l considérées. Or cela découle de ce que $\left|\int_{a+\varepsilon'}^{l}\right|$ ne surpasse pas

$$\overline{\underset{(a+\varepsilon', b)}{\text{borne}}}\ \varphi(x, n) \cdot \int_{a+\varepsilon'}^{b} |f(x) - f(a)|\,dx.$$

7. Corollaires. — Supposons que la fonction $\varphi(x, n)$ ayant $2b$ pour période est telle que $\varphi(x, n)$, de même que $\varphi(-x, n)$, sont, dans l'intervalle $(0, b)$, des noyaux positifs du premier ou second type (d'après la terminologie n° 3) avec la même valeur intégrale G. Si alors $f(x)$ est

une fonction périodique (à période $2b$) quelconque, sommable dans chaque intervalle fini, nous avons presque partout [4])

$$\int_0^{2b} f(x)\varphi(x-t,n)\,dx \to 2Gf(t).$$

8. Sur une propriété du noyau positif arbitraire. — Comme on sait, un point a s'appelle point de *densité nulle* (du coté droit) pour un ensemble mesurable e situé dans l'intervalle (a, b), si le rapport de la mesure de la partie de e appartenant à l'intervalle $(a, a+\delta)$, où $\delta > 0$, à la mesure de cet intervalle, savoir à δ, tend vers zéro avec δ. Le point a s'appelle point de *densité* 1 (du coté droit) pour l'ensemble mesurable e situé dans (a, b) s'il est un point de densité nulle pour l'ensemble complémentaire.

La définition ci-dessus donnée du point de densité nulle exige qu'un certain rapport tende vers zéro, mais aucune condition n'est imposée à la façon de laquelle cette tendance s'effectue. Introduisons à présent la définition suivante. Soit $\varepsilon(\delta)$ une grandeur positive déterminée pour toutes les valeurs positives suffisamment petites de δ et infiniment petite en même temps que δ [5]). Supposons que a soit un point de densité nulle (du coté droit) pour l'ensemble e situé dans (a, b). Nous dirons que cette densité nulle possède la puissance $\varepsilon(\delta)$ si le rapport infiniment petit qui figure dans la définition de la densité nulle est d'ordre supérieur à celui de $\varepsilon(\delta)$. Ensuite, si le point a est un point de densité 1 (du coté droit) pour l'ensemble e situé dans (a, b), nous dirons que cette densité 1 a la puissance $\varepsilon(\delta)$ si la densité nulle de l'ensemble complémentaire a la puissance $\varepsilon(\delta)$.

---

[4]) Pour voir cela il suffit de remarquer que quand $f(t)$ est sommable dans un certain intervalle (a,b), alors $\frac{1}{\varepsilon}\int_t^{t+\varepsilon}[f(x)-f(t)]\,dx$ et $\frac{1}{\varepsilon}\int_{t-\varepsilon}^{t}[f(x)-f(t)]\,dx$ tendent vers zéro avec ε presque partout dans cet intervalle (cette condition, en effet, est équivalente à la condition que $f(t)$ soit la dérivée de $\int_a^t f(x)\,dx$) et de plus $\frac{1}{2\varepsilon}\int_{t-\varepsilon}^{t+\varepsilon}|f(x)-f(t)|\,dx$ est infiniment petit avec ε presque partout dans cet intervalle (en effet, ce sera le cas en tous les points de continuité asymptotique de $f(t)$ si cette dernière est bornée; ensuite ça aura lieu en presque tous les points où $f(t)$, supposée sommable et à signe constant, est égale à zéro, car dans ce cas la condition considérée se réduit à la condition précédente; or, toute fonction sommable est la somme de trois fonctions: 1) bornée, 2) non négative et différente de zéro seulement sur un ensemble de mesure aussi petite que l'on veut, 3) non positive et différente de zéro seulement sur un ensemble de mesure aussi petite que l'on veut).

[5]) Nous supposons $\varepsilon(\delta)$ telle que $\underset{\delta \geqq \delta_0}{\text{borne}}\,\varepsilon(\delta) > 0$ pour toutes les valeurs de $\delta_0 > 0$. Ça aura lieu par exemple si $\varepsilon(\delta)$ est continue ou si elle ne croît pas quand δ tend vers zéro.

On sait ensuite qu'une fonction mesurable dans l'intervalle (a, b) s'appelle asymptotiquement continue au point a (du coté droit) s'il existe un ensemble mesurable situé dans (a, b), comprenant a et ayant a pour point de densité 1 (du coté droit), dans lequel la fonction donnée est continue au point a. Nous dirons que cette continuité asymptotique a la puissance $\varepsilon(\delta)$ si la densité 1 du dernier ensemble au point a possède la puissance $\varepsilon(\delta)$.

Ayant introduit ces définitions nous pouvons démontrer le théorème suivant.

Pour chaque noyau positif $\varphi(x, n)$ on peut trouver un tel $\varepsilon(\delta)$ dépendant de ce noyau seulement, que quelque soit la fonction $f(x)$ bornée et sommable dans (a, b) et asymptotiquement continue de puissance $\varepsilon(\delta)$ en a (du coté droit), nous ayons

$$\int_a^l f(x)\varphi(x, n)\,dx \to G f(a)$$

uniformément pour toutes les valeurs de l satisfaisant aux inégalités $a + \nu \leqq l \leqq b$, quelque petit que soit $\nu > 0$, où G est la valeur intégrale du noyau considéré.

Démontrons d'abord ce théorème pour un type particulier de noyaux, savoir pour les noyaux $\varphi(x, n)$ de propriété: quelque petit que soit $\nu > 0$, il existe un numéro N tel que, quand $n > N$, nous ayons $\varphi(x, n) = 0$ quand $a + \nu \leqq x \leqq b$. Soit $\varphi(x, n)$ un tel noyau. On peut alors trouver une suite de nombres positifs α_n tendant vers zéro de telle façon que $\varphi(x, n) = 0$ quand $a + \alpha_n \leqq x \leqq b$. Je dis que le théorème est valable si l'on prend

$$\varepsilon(\delta) = \min_{(1 \leqq n \leqq p(\delta))} \frac{1}{M_n},$$

où M_n est la borne supérieure de $\varphi(x, n)$ dans l'intervalle (a, b) et $p(\delta)$ le plus petit nombre naturel possédant la propriété que pour tout $n > p(\delta)$ nous ayons $\alpha_n < \delta$.

Soit e un ensemble quelconque pour lequel a est un point de densité nulle de puissance $\varepsilon(\delta)$. Pour $\delta < \delta_0$, où δ_0 est un nombre positif suffisamment petit, la densité de e dans $(a, a + \delta)$ sera inférieure à $\varepsilon(\delta)$. Montrons que, pour toutes les valeurs de n et pour $\delta < \delta_0$, nous avons

$$\int_{(a, a+\delta)e} \varphi(x, n)\,dx < \delta.$$

Soit d'abord $n > p(\delta)$. Alors $\alpha_n < \delta$ et l'intégrale considérée $\int_{(a, a+\delta)e}$ est égale à

$$\int_{(a, a+\alpha_n)e} \varphi(x, n)\,dx \leqq M_n\, m[(a, a + \alpha_n)e],$$

or la densité de e dans $(a, a+\alpha_n)$ est inférieure à $\varepsilon(\alpha_n)$, ensuite $\varepsilon(\alpha_n) \leqq \frac{1}{M_n}$ (car $n \leqq p(\alpha_n)$ parce que dans le cas contraire la supposition $n > p(\alpha_n)$ entrainerait, comme conséquence, $\alpha_n < \alpha_n$, ce qui est absurde); par conséquent

$$M_n m[(a, a+\alpha_n) e] \leqq M_n \alpha_n \varepsilon(\alpha_n) \leqq M_n \alpha_n \frac{1}{M_n} = \alpha_n < \delta.$$

Soit ensuite $n \leqq p(\delta)$. Alors la densité de e dans $(a, a+\delta)$

$$< \varepsilon(\delta) \leqq \frac{1}{M_n},$$

par conséquent l'intégrale considerée $\int_{(a, a+\delta) e}$ est inférieure à δ.

Soit, à présent, $f(x)$ une fonction bornée quelconque, sommable dans (a, b) asymptotiquement continue de puissance $\varepsilon(\delta)$ en a. Soit e un ensemble par rapport auquel elle est continue au point a et pour lequel a est un point de densité 1 de puissance $\varepsilon(\delta)$. Écrivons

$$\int_a^l f(x)\varphi(x, n)\,dx = f(a)\int_a^l \varphi(x, n)\,dx + \int_a^l [f(x) - f(a)]\varphi(x, n)\,dx.$$

Le premier terme tend vers $Gf(a)$ uniformément. Écrivons le second terme sous forme

$$\int_{a+\varepsilon'}^{l} + \int_{(a, a+\varepsilon') e} + \int_{(a, a+\varepsilon') ce}$$

où ε' est pris inférieur à ν et où ce désigne l'ensemble complémentaire à e. Eu égard à ce qui vient d'être démontré, $\int_{(a, a+\varepsilon') ce}$ est aussi petit que l'on veut, uniformément pour toutes les valeurs de n si ε' est suffisamment petit, car

$$\left| \int_{(a, a+\varepsilon') ce} \right| \leqq 2M \int_{(a, a+\varepsilon') ce} \varphi(x, n)\,dx,$$

où M est la borne supérieure de $|f(x)|$. Ensuite, en vertu de la continuité de $f(x)$ en a par rapport à l'ensemble e, $\int_{(a, a+\varepsilon') e}$ est aussi petit que l'on veut uniformément pour toutes les valeurs de n si ε' est suffisamment petit, car

$$\left| \int_{(a, a+\varepsilon') e} \right| \leqq \overline{\text{borne}}_{(x \text{ dans } (a, a+\varepsilon') e)} |f(x) - f(a)| \cdot N,$$

où N est un nombre supérieur à toutes les $\int_a^b \varphi(x, n)\,dx$. Enfin $\int_{a+\varepsilon'}^{l}$ tend vers zéro uniformément pour les valeurs considérées de l.

Soit, à présent, $\varphi(x, n)$ un noyau positif arbitraire. On peut choisir des nombres positifs α_n de telle sorte qu'ils tendent vers zéro, mais si

lentement que $\int_{a+\alpha_n}^{b} \varphi(x,n)\,dx$ tende vers zéro. Posons

$$\varphi_1(x,n) = \begin{cases} \varphi(x,n) & \text{pour } a \leqq x \leqq a+\alpha_n; \\ 0 & \text{pour } a+\alpha_n < x \leqq b. \end{cases}$$

Il est évident que $\varphi_1(x,n)$ est un noyau à valeur intégrale G, par conséquent $\varphi(x,n) - \varphi_1(x,n)$ est un noyau à valeur intégrale 0. De plus $\varphi(x,n) - \varphi_1(x,n) \geqq 0$. Pour démontrer complètement notre théorème, dont la valadité pour le noyau $\varphi_1(x,n)$ est déjà prouvée, il suffit seulement de voir que

$$\int_a^l f(x)\,[\varphi(x,n) - \varphi_1(x,n)]\,dx \to 0$$

uniformément pour toutes les valeurs de l qui satisfont aux inégalités $a+\nu \leqq l \leqq b$. Or il en est precisément ainsi, car la valeur absolue de la dernière intégrale est

$$\leqq M \int_a^b [\varphi(x,n) - \varphi_1(x,n)]\,dx,$$

où M est la borne supérieure de $|f(x)|$.

9. Remarques sur la propriété fondamentale du noyau arbitraire. — Ainsi que l'on sait[6]), la propriété fondamentale du noyau arbitraire consiste en ce que si $\varphi(x,n)$ est un noyau quelconque (à valeur intégrale G) qui satisfait à la seule condition que les nombres $\int_\alpha^\beta \varphi(x,n)\,dx$ (pour tous les nombres naturels n et pour tous les α, β appartenant à l'intervalle a, b) soient bornés dans leur ensemble et si $f(x)$ est une fonction quelconque à variation bornée (dans (a,b)), alors $\varphi(x,n)f(x)$ est un noyau à valeur intégrale $Gf(a+0)$.

Le fait que les nombres $\int_\alpha^\beta \varphi(x,n)\,dx$ sont bornés dans leur ensemble dans le cas de noyau à signe constant est une conséquence directe de la définition du noyau; mais dans le cas de noyau à signe variable il n'est pas nécéssaire. Considérons un exemple d'un noyau $\varphi(x,n)$ pour lequel les nombres $\int_\alpha^\beta \varphi(x,n)\,dx$ ne sont pas bornés dans leur ensemble. Posons

$$\varphi(x,n) = \begin{cases} K_n \sin\frac{2\pi(x-a)}{\varepsilon_n} & \text{pour } a \leqq x \leqq a+\varepsilon_n; \\ 0 & \text{pour } a+\varepsilon_n < x \leqq b; \end{cases}$$

[6]) Voir par exemple E. Picard, Traite d'Analyse, vol. I, troisième édition (Paris, 1922), p. 274.

où K_n et ε_n sont de nombres positifs et $\varepsilon_n \to 0$. Il est clair que $\varphi(x, n)$ est un noyau à valeur intégrale 0. Or

$$\int_a^{a+\frac{\varepsilon_n}{2}} \varphi(x, n)\, dx = \frac{K_n \varepsilon_n}{\pi},$$

par conséquent si $K_n \varepsilon_n$ ne sont pas bornés $\left(\text{par exemple si } K_n = \frac{1}{\varepsilon_n^2}\right)$, $\int_\alpha^\beta \varphi(x, n)\, dx$ ne sont pas bornés dans leur ensemble. Pour un tel noyau la propiété fondamentale enoncée peut ne pas avoir lieu. Prenons, par exemple, $f(x) = x - a$. Alors

$$\int_a^b f(x)\, \varphi(x, n)\, dx$$

$$= \int_a^{a+\varepsilon_n} (x - a)\, K_n \sin \frac{2\pi(x-a)}{\varepsilon_n}\, dx = \frac{K_n \varepsilon_n^2}{4\pi^2} \int_0^{2\pi} t \sin t\, dt = -\frac{K_n \varepsilon_n^2}{2\pi},$$

ce qui ne tend pas vers zéro si $K_n \varepsilon_n^2$ ne tend pas vers zéro $\left(\text{par exemple si } K_n = \frac{1}{\varepsilon_n^2}\right)$.

Considérons encore un exemple de noyau $\varphi(x, n)$ avec $\int_\alpha^\beta \varphi(x, n)\, dx$ bornés dans leur ensemble, mais tel que, dans chaque intervalle (α, β) situé dans (a, b), $\varphi(x, n)$ ne soit bornée ni inférieurement ni supérieurement (pour l'ensemble de tous les n). Posons

$$\varphi(x, n) = (-1)^i K_n \quad \text{pour} \quad \frac{i-1}{2n} < x < \frac{i}{2n} \qquad (i = 1, 2, \ldots, 2n)$$

$\left(\text{aux points } \frac{i}{2n} \text{ on peut poser } \varphi(x, n) = 0\right)$. Alors $\varphi(x, n)$ sera déterminée dans $(0, 1)$. Si l'on prend les nombres K_n de manière que $K_n \to \infty$, mais $\frac{K_n}{n} \to 0$ (par exemple $K_n = \sqrt{n}$), on voit aisément que $\varphi(x, n)$ est un noyau à valeur intégrale 0, avec $\int_\alpha^\beta \varphi(x, n)\, dx$ bornés dans leur ensemble, et non borné (pour l'ensemble de tous les n) ni inférieurement ni supérieurement dans aucun intervalle situé dans (a, b).

10. Sur l'intégrale $\int_a^b f(x)\, \varphi(nx)\, dx$, où $\varphi(x)$ est une fonction périodique. — Si la fonction $\varphi(x)$ (que nous supposons bornée, $|\varphi| \leqq M$, et sommable dans chaque intervalle fini) a la période ω et $\int_0^\omega \varphi(x)\, dx = 0$, alors pour chaque fonction $f(x)$, sommable dans l'intervalle (a, b), et pour n croissant indéfiniment, nous avons

$$\int_a^b f(x)\, \varphi(nx)\, dx \to 0,$$

et plus encore même,

$$\int_{\alpha}^{\beta} f(x)\,\varphi(nx)\,dx \to 0$$

uniformément pour toutes les valeurs de α et β situés dans l'intervalle (a, b) [7]).

11. Le noyau $\frac{\varphi(nx)}{x}$, où φ est une fonction périodique [8]). — Soit $\varphi(x)$ une fonction périodique possédant les propriétés: elle est bornée, elle est sommable dans chaque intervalle fini, elle a la période ω, $\int_0^{\omega} \varphi(x)\,dx = 0$, $\varphi(0) = 0$, ses nombres dérivés au point $x = 0$ sont finis. Alors

$$\frac{\varphi(nx)}{x}$$

est un noyau à valeur intégrale

$$G = \int_0^{+\infty} \frac{\varphi(x)}{x}\,dx \quad [9]).$$

---

[7]) Dans le cas où $f(x)$ est une fonction continue, cette proposition se démontre à peu près comme dans la théorie classique des séries trigonométriques on démontre que les coéfficients de Fourier tendent vers zéro: Soit $\varepsilon(\delta)$ le module de continuité de la fonction $f(x)$. De

$$\int_{\alpha}^{\beta} f(x)\,\varphi(nx)\,dx = \int_{\alpha}^{\alpha+\frac{p\omega}{n}} + \int_{\alpha+\frac{p\omega}{n}}^{\beta}$$

$$= \sum_{i=1}^{p} \int_{\alpha+\frac{(i-1)\omega}{n}}^{\alpha+\frac{i\omega}{n}} \left[f(x) - f\left(\alpha+\frac{(i-1)\omega}{n}\right)\right]\varphi(nx)\,dx + \int_{\alpha+\frac{p\omega}{n}}^{\beta} f(x)\,\varphi(nx)\,dx,$$

où l'entier p satisfait aux inégalités

$$\alpha+\frac{p\omega}{n} \leqq \beta < \alpha+\frac{(p+1)\omega}{n},$$

il résulte que

$$\left|\int_{\alpha}^{\beta} f(x)\,\varphi(nx)\,dx\right| \leqq (b-a)\,M\,\varepsilon\left(\frac{\omega}{n}\right) + \frac{\omega}{n}\,M \max|f|.$$

Si $f(x)$ est seulement sommable, alors elle est $= f_1 + f_2$, où f_1 est continue et $\int_a^b |f_2|\,dx < \varepsilon$ (ε étant donné à l'avance). Alors

$$\int_{\alpha}^{\beta} f(x)\,\varphi(nx)\,dx = \text{inf. pet. unif.} + \int_{\alpha}^{\beta} f_2(x)\,\varphi(nx)\,dx,$$

or la valeur absolue du dernier terme est $\leqq \varepsilon M$.

[8]) Les choses exposées dans ce n° peuvent être regardées comme une certaine généralisation de certains faits bien connus de la théorie classique des séries trigonométriques.

[9]) La dernière intégrale est convergente (à cause de ce que $\frac{\varphi(x)}{x}$ est borné près de l'origine et en vertu du critère d'Abel de convergence des intégrales); par consé-

(Voir la fin de cette note [9]) à la page suivante.)

Le noyau $\frac{\varphi(nx)}{x}$ possède la propriété suivante.

Si la fonction $f(x)$, sommable dans $(0, b)$, est telle que $\frac{f(x)-f(0)}{x}$ est sommable près de l'origine (par exemple $f(x)$ satisfait à la condition de Lipschitz d'ordre quelconque pour $x=0$), alors

$$\int_0^l f(x)\frac{\varphi(nx)}{x}\,dx \to G f(0)$$

uniformément pour toutes les valeurs de l satisfaisant aux inégalités $\nu \leqq l \leqq b$, quelque petit que soit $\nu > 0$, et où $G = \int_0^{+\infty} \frac{\varphi(x)}{x}\,dx$.

Pour le démontrer écrivons

$$\int_0^l f(x)\frac{\varphi(nx)}{x}\,dx = f(0)\int_0^l \frac{\varphi(nx)}{x}\,dx + \int_0^l \frac{f(x)-f(0)}{x}\varphi(nx)\,dx,$$

et à présent il suffit seulement de se reporter au théorème n° 10 pour en conclure que le second terme tend vers zéro uniformément.

Nous en tirons comme conséquence:

Si $f(x)$ est sommable dans $(0, b)$, alors

$$\int_0^b f(x)\frac{\varphi(n(x-t))}{x-t}\,dx \to 2Gf(t), \quad \text{où} \quad 2G = \int_{-\infty}^{+\infty}\frac{\varphi(x)}{x}\,dx,$$

pour toutes les valeurs de t satisfaisant aux inégalités $0 < t < b$, autour desquelles $\frac{f(x)-f(t)}{x-t}$ est sommable (par exemple $f(x)$ satisfait à la condition de Lipschitz d'ordre quelconque pour $x=t$).

---

quent pour tout $\varepsilon > 0$ on peut trouver un K tel que pour $k > K$ $\int_0^k \frac{\varphi(x)}{x}\,dx$ diffère de $\int_0^{+\infty}\frac{\varphi(x)}{x}\,dx$ de moins qu'ε. Or à présent

$$\int_0^l \frac{\varphi(nx)}{x}\,dx = \int_0^{nl}\frac{\varphi(x)}{x}\,dx,$$

par conséquent pour $n > \frac{K}{\nu}$ nous avons

$$\left| G - \int_0^l \frac{\varphi(nx)}{x}\,dx \right| < \varepsilon$$

pour tous les l satisfaisant à l'inégalité $\nu \leqq l$.

Soit maintenant $\psi(x)$ une fonction à variation bornée dans $(0, b)$. Alors (en vertu de la propriété fondamentale des noyaux) $\frac{\varphi(nx)}{x}\psi(x)$ est un noyau à valeur intégrale $G\psi(0+0)$, où $G = \int_0^{+\infty} \frac{\varphi(x)}{x} dx$. Absolument comme ci dessus il résulte que si la fonction $f(x)$ est sommable dans $(0, b)$ et le rapport $\frac{f(x)-f(0)}{x}$ est sommable près de l'origine (par exemple $f(x)$ satisfait à la condition de Lipschitz d'ordre quelconque en $x = 0$), alors

$$\int_0^l f(x)\frac{\varphi(xn)}{x}\psi(x)\,dx \to G\psi(0+0)f(0)$$

uniformément pour toutes les valeurs de l qui satisfont aux inégalités $\nu \leqq l \leqq b$, quelque petit que soit $\nu > 0$.

Il suit de là que si $\psi(x)$ est une fonction périodique (à periode ω) qui s'annule pour les valeurs égales aux multiples de ω et qui a la borne inférieure du module différente de zéro dans tout intervalle qui ne contient pas de multiples de ω, qui a une variation bornée dans chaque intervalle fini et, enfin, telle que $\frac{x}{\psi(x)}$ soit à variation bornée autour de l'origine et tend vers l'unité quand x tend vers zero, alors

$$\int_0^{\omega} f(x)\frac{\varphi(n(x-t))}{\psi(x-t)}dx \to 2Gf(t), \qquad \text{où} \quad 2G = \int_{-\infty}^{+\infty} \frac{\varphi(x)}{x}dx,$$

pour toutes les valeurs de t autour desquelles $\frac{f(x)-f(t)}{x-t}$ est sommable, $f(x)$ étant une fonction périodique (à période ω) sommable (dans chaque intervalle fini).

Moscou, Mars 1924.

(Eingegangen am 12. September 1929.)

Die numerische Lösung der Gleichungen vierten Grades.

Von

Friedrich Schilling in Danzig-Langfuhr.

Es sei die Gleichung vierten Grades mit der Unbekannten t in der reduzierten Form gegeben:

$$t^4 + \lambda \cdot t^2 + \mu \cdot t + \nu = 0 \tag{1}$$

mit den reellen (endlichen) Koeffizienten λ, μ, ν. Wir wollen untersuchen: *Wie viele reelle Wurzeln ergeben sich für gegebene Werte λ, μ, ν und wie lassen sich diese Wurzeln leicht bestimmen?*

Wir schreiben diese Gleichung in der Form:

$$t^4 + (1 - 2p) \cdot t^2 - 2q \cdot t + (p^2 + q^2 - \varrho^2) = 0, \tag{2}$$

so daß also

$$\begin{cases} \lambda = 1 - 2p, \\ \mu = -2q, \\ \nu = p^2 + q^2 - \varrho^2 = \frac{(1-\lambda)^2 + \mu^2}{4} - \varrho^2 \end{cases} \tag{3a, b, c}$$

gesetzt ist. Es sind hiernach die Größen p, q, ϱ^2 eineindeutig den Größen λ, μ, ν zugeordnet, wobei für reelle Werte λ, μ, ν sehr wohl ϱ selbst rein imaginär sein kann. Neben die Gleichung (2) stellen wir jetzt das Gleichungspaar mit den Unbekannten x, y:

$$\begin{cases} y^2 = x, \\ (x - p)^2 + (y - q)^2 = \varrho^2. \end{cases} \tag{4a, b}$$

Es folgt sogleich der Satz:

1. Die aus den beiden Gleichungen (4a, b) folgenden Wurzeln für y sind zugleich die Wurzeln der Gleichung (2).

Wir bemerken ferner:

2. Wenn ϱ^2 negativ, d. h. $\nu > p^2 + q^2$ oder $\nu > \frac{(1-\lambda)^2 + \mu^2}{4}$ ist, so ergeben sich keine reellen Wurzeln, so daß dieser Fall für das Folgende schon ausgeschieden werden kann.

3. Man zeichne die Parabel mit der Gleichung (4a) auf Koordinatenpapier mit den Achsen x, y und ferner für die gegebenen Werte p, q, ϱ den Kreis mit der Gleichung (4b) um den Mittelpunkt $(x, y) = (p, q)$ mit dem (reellen) Radius ϱ. Die Schnittpunkte beider Kurven ergeben als ihre y-Koordinaten die gesuchten Wurzeln t.

Die Ausgestaltung dieser Konstruktion zu einem beweglichen Nomogramm für die verschiedenen Werte p, q, ϱ habe ich in der besonderen Arbeit[1]): „Über einige neue Nomogrammarten" ausgeführt.

In der vorliegenden Arbeit soll es sich jedoch, wie gesagt, darum handeln, nun aus der einfachen geometrischen Konstruktion die Kriterien dafür zu gewinnen, wann die gegebene Gleichung (1) oder (2) keine, zwei oder vier reelle und voneinander verschiedene Wurzeln besitzt, sowie auch insbesondere wann eine oder zwei reelle Doppelwurzeln oder eine dreifache oder eine vierfache Wurzel sich ergeben.

Ersichtlich kommt diese Frage geometrisch darauf hinaus, zu bestimmen, wann der Kreis (4b) die Parabel (4a) gar nicht, zweimal oder viermal schneidet bzw. einmal oder zweimal berührt oder oskuliert oder in vier zusammenfallenden Punkten berührt.

Dem Satze 3 gemäß stellt uns die ∞^3-fache Gesamtheit aller Kreise der (x, y)-Ebene gleichsam die (zyklographische) Mannigfaltigkeit dar, in der wir uns bewegen. Wir können aber auch die Größen p, q, ϱ als rechtwinklige (x, y, z)-Koordinaten des Raumes deuten, indem wir noch die z-Achse hinzunehmen. Dem Inbegriff der Werte p, q, ϱ^2 entsprechen dann also zwei zur (x, y)-Ebene symmetrische Punkte des (x, y, z)-Raumes.

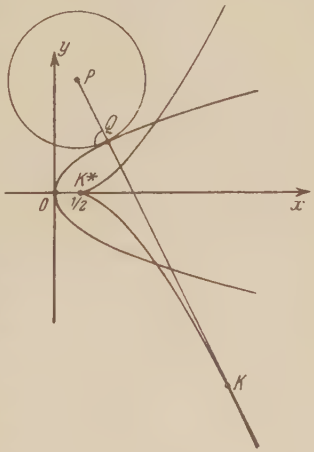


Fig. 1.

Wir erkennen sogleich: Wenn wir in der Fig. 1 den Punkt P die Normale eines Punktes Q der Parabel durchlaufen lassen, so bekommen wir als Darstellung der Werte p, q, ϱ für den Fall der Doppelwurzel in der Kreismannigfaltigkeit die ∞^1-Kreise, welche die Parabel im Punkte Q berühren, im (x, y, z)-Raum aber als Ort der Punkte (x, y, z) oder (p, q, ϱ) die Punkte zweier Geraden, welche in der zur (x, y)-Ebene senkrechten Ebene durch die Normale liegen und zu der Normalen im Punkte Q unter dem Winkel 45^0 geneigt sind. (Ist insbesondere P im Punkte Q selbst gewählt, so ist ja

[1]) Zeitschrift f. mathematischen u. naturwissenschaftlichen Unterricht, 61. Jahrg. (1930), S. 241ff. In jener Arbeit sind aus Zweckmäßigkeitsgründen die Koordinaten x, y im Vergleich mit der vorliegenden Arbeit vertauscht.

$\varrho = 0$, also $\nu = p^2 + q^2 = \frac{(1-\lambda)^2+\mu^2}{4}$; in allen anderen Fällen ist $\nu < \frac{(1-\lambda)^2+\mu^2}{4}$.) Wir denken nun die Evolute zu der Parabel hinzugezeichnet (Fig. 1), also die Neilsche oder semikubische Parabel mit der Gleichung

$$(5) \qquad \left(x - \frac{1}{2}\right)^3 = \frac{27}{16} y^2,$$

oder mit den Gleichungen mit dem Parameter τ:

$$(6a, b) \qquad \begin{cases} x = 3\tau^2 + \frac{1}{2}, \\ y = 4\tau^3. \end{cases}$$

Wir wollen weiter gleich darauf hinweisen, daß in dem Falle, wo der Punkt P mit dem Berührungspunkt K der Normalen und der Evolute, dem Krümmungsmittelpunkt der Parabel für den Punkt P, zusammenfällt, dann natürlich der Fall des Kreises mit dem Radius KQ einer dreifachen Wurzel entspricht. Ist der Punkt K insbesondere die Spitze K^* der Evolute, so entspricht dem Kreise um K^* durch den Scheitelpunkt O der Parabel eine vierfache Wurzel (Fig. 1).

Wir gewinnen ferner sogleich den Satz:

4. Der geometrische Ort aller Raumpunkte (x, y, z) oder (p, q, ϱ), denen eine Doppelwurzel der Gleichung (2) entspricht, d. h. die besondere Art einer Diskriminantenfläche der Gleichung (2), ist die Böschungsfläche der (horizontal gelegenen) Parabel $y^2 = x$ für den Böschungswinkel 45^0.

Die Böschungsflächen der Parabel habe ich in der Arbeit „Über die Böschungsflächen mit Kegelschnitten als Basiskurven“ in der Zeitschrift für angewandte Mathematik und Mechanik **3** (1923), S. 197—217 behandelt, worauf wegen aller Einzelheiten verwiesen sei. Für den Parameter $\frac{1}{2}$ unserer Parabel $y^2 = x$ und für den Böschungswinkel $\delta = 45^0$ lautet dann die Gleichung unserer Böschungsfläche gemäß der Gleichung (50′) der Seite 213 daselbst:

$$(7) \quad \left(\frac{1}{2} + 2x\right)^2 \cdot (-x + y^2 - z^2)^2 - 2z^2\left(\frac{1}{2} + 2x\right)^3 - \frac{27}{4} z^4 - 9z^2\left(\frac{1}{2} + 2x\right)\cdot(-x + y^2 - z^2) + 4(-x + y^2 - z^2)^3 = 0.^2)$$

5. Die Böschungsfläche ist also eine Fläche 6. Ordnung, die zu den (x, y)- und (x, z)-Ebenen symmetrisch ist.

Da die Böschungsfläche auch von allen Ebenen eingehüllt wird, deren einzelne durch je eine Tangente der Parabel $y^2 = x$ geht und gegen die

---

[2]) Die analoge Gleichung der Böschungsfläche der Ellipse ist zuerst von Catalan und Cayley aufgestellt (vgl. G. Loria, Spezielle algebraische und transzendente ebene Kurven, Bd. II, 2. Aufl., Leipzig 1911, S. 282).

(x, y)-Ebene unter 45^0 geneigt ist, so ist sie eine abwickelbare Fläche. Ihre Rückkehrkante hat die Gleichungen mit dem Parameter τ:

$$(8\,\mathrm{a, b, c}) \qquad \begin{cases} x = 3\tau^2 + \frac{1}{2}, \\ y = 4\tau^3, \\ z = \pm \frac{1}{2}(4\tau^2 + 1)^{\frac{3}{2}}. \end{cases}$$

Auf die (x, y)-Ebene projiziert ergibt diese Rückkehrkante eben die Evolute der Parabel mit den Gleichungen (6a, b).

Wir wollen nun zum Vergleich einen Blick auf die Art der Diskriminantenfläche werfen, wie sie von F. Klein[3]) angegeben ist. Wenn wir mit Klein jetzt die Koeffizienten λ, μ, ν der Gleichung (1) mit X, Y, Z bezeichnen, so ist diese Art der Diskriminantenfläche durch die Gleichung

$$\left(Z + \frac{X^2}{12}\right)^3 - 27\left(\frac{XZ}{6} - \frac{Y^2}{16} - \frac{X^3}{216}\right)^2 = 0$$

in den rechtwinkligen (X, Y, Z)-Koordinaten bestimmt. Diese Gleichung, in der die Glieder 6. Ordnung sich aufheben, stellt eine Fläche 5. Ordnung dar. Durch die quadratische Transformation der Gleichungen (3a, b, c), in der ja die Größen λ, μ, ν durch X, Y, Z und die Größen p, q, ϱ durch die Koordinaten x, y, z dargestellt sind, geht diese Diskriminantenfläche von Klein in unsere Art der Diskriminantenfläche mit der Gleichung (7) über.

Von unserer Böschungsfläche können wir uns sehr leicht eine anschauliche Vorstellung schaffen. Die horizontalen Schnitte der Fläche (bei vertikal gedachter z-Achse) sind auf die (x, y)-Ebene projiziert einfach die einzelnen aus zwei Zweigen bestehenden Äquidistanten der Parabel $y^2 = x$ und zwar für einen Abstand, der durch die z-Koordinate des Schnittes gegeben ist. In der Fig. 2, welche einen Teil der oberen Hälfte der Fläche darstellt, erkennen wir als gewählte obere Begrenzung der Fläche eine solche Äquidistante. Insbesondere sei nun noch das Resultat hervorgehoben, das eben zu einem analogen Resultat der Darstellung von Klein in engster Beziehung steht:

Durch diese Diskriminantenfläche, und zwar durch ihre eine Hälfte für $z \geqq 0$, auf die wir uns beschränken können, zerfällt der Halbraum in drei Gebiete (vgl. die Fig. 2). Das erste dieser Gebiete besteht aus den beiden Teilen Ia und Ib, die sich unmittelbar über der Ebene $z = 0$ und

[3]) F. Klein, Elementarmathematik vom höheren Standpunkte aus, Bd. I, Berlin 1924, S. 103—109. Daselbst ist auch auf das Hartensteinsche Fadenmodell dieser Diskriminantenfläche, dessen Herausgabe von mir geleitet wurde, hingewiesen, das im Verlage von M. Schilling in Leipzig erschienen ist. Vgl. auch die dazugehörige Arbeit: R. Hartenstein, Die Diskriminantenfläche der Gleichungen vierten Grades, Leipzig 1909.

zwar über dem Äußeren und dem Inneren der Parabel $y^2 = x$ erheben, und zwar wird der Teil Ib noch durch die Doppelkurve der Fläche in der (x, z)-Ebene, die Parabel $z^2 = x - \frac{1}{4}$, begrenzt, soweit diese nicht isoliert verläuft. Das Gebiet II liegt mit seinen Horizontalschnitten zwischen den beiden Zweigen der Äquidistanten. Das noch übrige Gebiet III aber wird in seinen Horizontalschnitten für $z \geqq \frac{1}{2}$ von den Teilen der inneren Äquidistantenzweige begrenzt, die aus den drei Bogen zwischen den beiden Spitzen und dem Doppelpunkt bestehen. Es gilt dann der Satz:

6a. Liegt der Punkt P, dessen Koordinaten die Größen p, q, ϱ darstellen bzw. in dem Gebiete I, II oder III, so hat die Gleichung (1) ent-

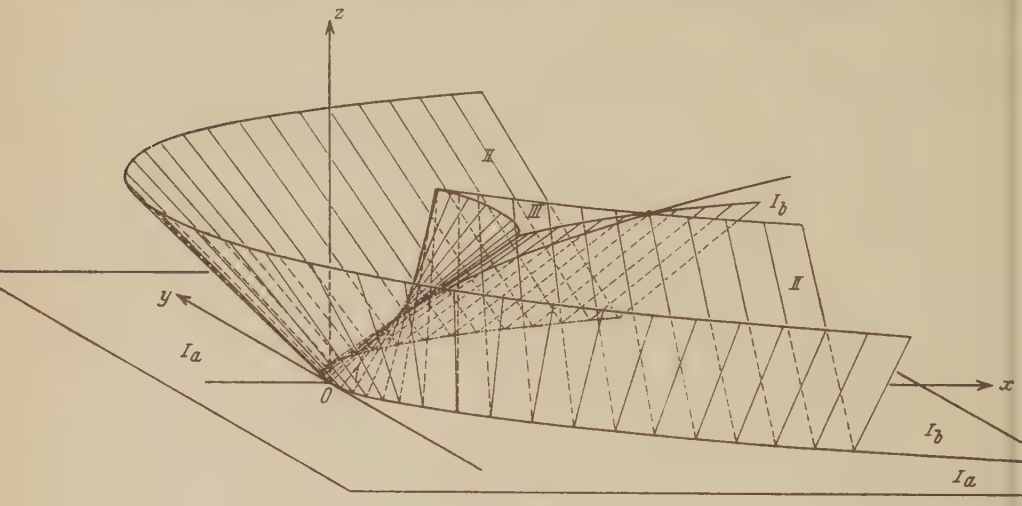


Fig. 2.

sprechend keine, zwei oder vier reelle Wurzeln (vgl. indessen auch zur Ergänzung den Satz 2).

6b. Liegt aber der Punkt P auf der Doppelkurve in der (x, z)-Ebene, auf der Parabel $z^2 = x - \frac{1}{4}$, so hat die Gleichung (1) zwei Doppelwurzeln, die zusammenfallen, wenn der Punkt P gerade in der Spitze der Rückkehrkurve unserer Fläche liegt (und beide konjugiert-imaginär sind, wenn der Punkt P auf dem isolierten Teil der Parabel liegt).

6c. Liegt der Punkt P auf der Rückkehrkante, so hat die Gleichung (1) eine dreifache Wurzel und noch eine einfache Wurzel, und liegt endlich der Punkt P auf einem der noch übrigen Teile der Diskriminantenfläche, so hat die Gleichung (1) eine Doppelwurzel und außerdem noch zwei weitere reelle Wurzeln oder bzw. nicht, je nachdem der Teil der Diskriminantenfläche die Raumgebiete II und III oder bzw. die Raumgebiete Ia und II oder Ib und II oder Ia und Ib voneinander trennt.

Diese Resultate für die verschiedenen Möglichkeiten der Wurzeln der Gleichung (1) sind hiernach auch leicht rein analytisch auszusprechen.

Nun aber liegt der wesentliche Vorteil der von uns entwickelten Methode gar nicht darin, daß wir diese neue, allerdings sehr anschauliche Art der Diskriminantenfläche entwickelt haben. Wir können vielmehr jetzt die Diskriminantenfläche in die zyklographische Mannigfaltigkeit der (x, y)-Ebene abbilden. Und dieses ebene Abbild ist den räumlichen Darstellungen weit überlegen. Diese Überlegenheit wirkt sich sogar noch weitergehend dahin aus, daß wir in dieser ebenen Darstellung auch für jedes Wertetripel λ, μ, ν oder p, q, ϱ die reellen Wurzeln sofort in einfachster Weise graphisch ablesen können. Im einzelnen sei hierzu folgendes bemerkt: Durch die Parabel, ihre Evolute und den Teil $x \geqq \frac{1}{2}$ der x-Achse wird die ganze (x, y)-Ebene jetzt in die Gebiete I II, IIIa, b, IVa, b zerlegt, wie die Fig. 3 es zeigt. Die Gebiete I, II zusammen bilden das „Innere", die Gebiete IIIa, b und IVa, b das „Äußere" der Evolute. Wir wissen:

7. Zu jedem Punkte (x, y) oder (p, q) der Gebiete I, II gehört nur ein Kreis, der die Parabel berührt (Fall 1), zu jedem Punkte (x, y) oder (p, q) der Gebiete IIIa, b und IVa, b aber drei solche Kreise (Fall 2).

Wird jetzt im ersten Falle der Radius des berührenden Kreises mit ϱ_1 bezeichnet, so ergibt sich weiter der Satz:

8a. Ist für den Fall, daß der Punkt (p, q) dem Gebiete I oder II angehört, der zugehörige Wert ϱ gemäß der gegebenen Gleichung (2) kleiner oder größer als ϱ_1 oder gleich ϱ_1, so hat diese Gleichung insgesamt bzw. keine oder zwei reelle Wurzeln oder eine Doppelwurzel (Fig. 3).

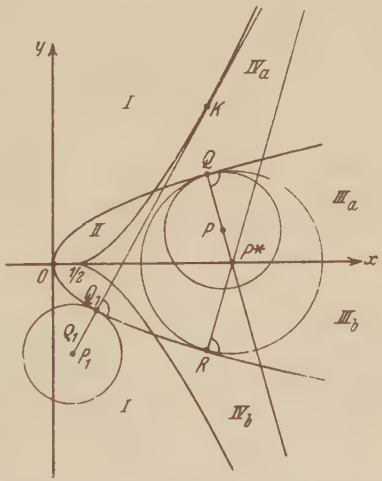


Fig. 3.

Im zweiten Falle wollen wir jedoch zunächst noch folgendes erkennen: Die Mittelpunkte der Kreise, welche die Parabel doppelt berühren, liegen auf dem Teil $x \geqq \frac{1}{2}$ der x-Achse. Ist nun P^* ein beliebiger solcher Punkt mit den zugehörigen Berührungspunkten Q, R des doppelt berührenden Kreises, so berührt ersichtlich jeder Kreis um einen Mittelpunkt P auf P^*Q oder P^*R mit dem Radius PQ bzw. PR die Parabel im Punkte Q bzw. R, ohne sie jedoch noch weiter zu schneiden. Ist daher P z. B. ein beliebiger Punkt des Bereiches IIIa (also mit positiver Ordinate), so schneidet der kleinste Kreis, der die Parabel berührt, diese nicht weiter, und die Gerade PQ seines

Berührungsradius ist eine Tangente an den Zweig der Evolute für $y < 0$. Analoges gilt für einen beliebigen Punkt P des Bereiches IVa. Die den beiden anderen berührenden Kreisen zugehörigen Geraden der Berührungsradien gehören dagegen als Tangenten stets dem Zweige der Evolute für $y > 0$ an.

Es sei nun ein beliebiger Punkt P der Bereiche IIIa, b oder IVa, b gegeben; die drei die Parabel berührenden Kreise um P mögen die Radien $\varrho_1 < \varrho_2 < \varrho_3$ haben. Es folgt dann also der Satz:

8b. Gilt, im Falle der Punkt (p, q) dem Gebiet III oder IV angehört (Fig. 4a, b), für die zugehörigen Werte ϱ

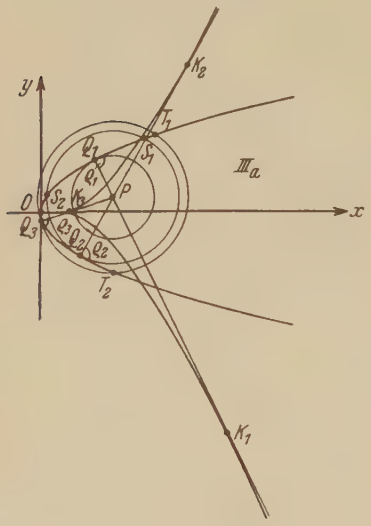


Fig. 4a.

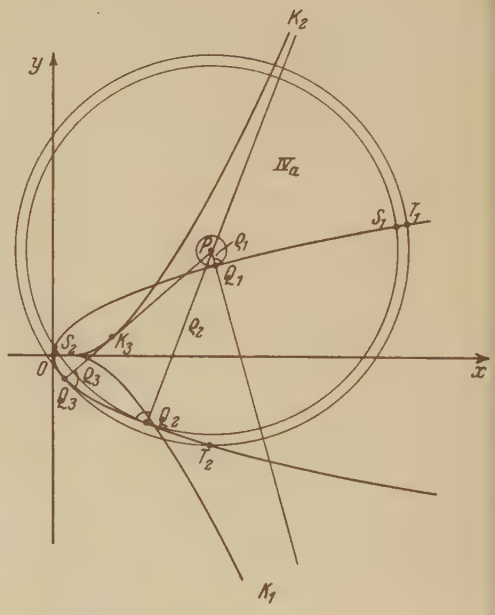


Fig. 4b.

bzw. $\varrho < \varrho_1$, $\varrho = \varrho_1$, $\varrho_1 < \varrho < \varrho_2$, $\varrho = \varrho_2$, $\varrho_2 < \varrho < \varrho_3$, $\varrho = \varrho_3$, $\varrho > \varrho_3$, so hat die Gleichung (1) insgesamt bzw. keine reelle Wurzel, eine Doppelwurzel, zwei reelle Wurzeln, zwei reelle Wurzeln und eine Doppelwurzel, vier reelle Wurzeln, zwei reelle Wurzeln und eine Doppelwurzel, zwei reelle Wurzeln.

Diese Resultate sind dann noch leicht für die Fälle zu ergänzen, daß der Punkt P auf einer der Trennungslinien der Gebiete liegt.

Endlich sei noch einmal dem Satz 1 gemäß hervorgehoben:

9. Für jede Gleichung (1) oder (2) ergeben die Ordinaten der Schnittpunkte der Parabel $y^2 = x$ und des Kreises um den Punkt $(x, y) = (p, q)$ mit dem Radius ϱ die reellen Wurzeln.

Wir wollen noch die Bemerkung hinzufügen:

10. In der vorstehend beschriebenen Abbildung der Diskriminantenfläche mit der Gleichung (7) in die zyklographische Mannigfaltigkeit der (x, y)-Ebene ist ein interessantes und lehrreiches Beispiel für die von E. Müller in Wien begründeten neuen Abbildungsmethoden in der Darstellenden Geometrie gegeben [4]).

[4]) Vgl. in dieser Hinsicht z. B. L. Eckhart, Konstruktive Abbildungsverfahren, eine Einführung in die neueren Methoden der Darstellenden Geometrie, Wien 1926, S. 77 ff.

(Eingegangen am 9. April 1930.)

Projektive Invarianten einer Kurvenkongruenz und einer Kurve.

Von

V. Hlavatý in Prag.

Es sei eine n-dimensionale Mannigfaltigkeit X_n gegeben, die auf die Koordinaten x^k [1]) bezogen ist. Unter einer Konnexion in X_n verstehen wir den Inbegriff von $n^2(n+1)/2$ Funktionen $\Gamma^k_{ij} = \Gamma^k_{ji}$ von x, die sich beim Übergang zu einem anderen Koordinatensystem x^ν [2]) in $\Gamma^\nu_{\lambda\mu}$ transformieren:

$$\Gamma^k_{ij} = A^k_\nu A^\lambda_i A^\mu_j \Gamma^\nu_{\lambda\mu} + A^k_\lambda A^\mu_j \partial_\mu A^\lambda_i . \tag{1}$$

Dabei ist

$$A^k_\lambda = \frac{\partial x^k}{\partial x^\lambda}, \qquad A^\nu_k = \frac{\partial x^\nu}{\partial x^k}, \qquad \partial_\mu = \frac{\partial}{\partial x^\mu} .$$

Mittels der Konnexion Γ können wir die *geodätischen* Kurven konstruieren:

$$\frac{dx^k}{dt} + \Gamma^k_{ij} \frac{dx^i}{dt} \frac{dx^j}{dt} = 0 . \tag{2}$$

Die Konnexion

$$'\Gamma^k_{ij} = \Gamma^k_{ij} + A^k_i p_j + A^k_j p_k \qquad (A^k_i = A^k_\lambda A^\lambda_i = \delta^k_i) \tag{3}$$

mit dem beliebigen kovarianten Vektor p_k ist die allgemeinste symmetrische (d. h. mit $'\Gamma^k_{ij} = {}'\Gamma^k_{ji}$), welche dieselben geodätischen Linien hat wie Γ [3]). Aus diesem Grunde wollen wir (3) eine „bahntreue Transformation" der Konnexion nennen [4]). Durch Elimination von p_i aus (3) kommt man zu

---

[1]) In der ganzen Arbeit durchlaufen die Indizes i, j, k, l die n Symbole $1 \ldots n$, während a, b, c, d außer $1 \ldots n$ noch das $(n+1)$-te Symbol 0 durchzulaufen haben.

[2]) In der ganzen Arbeit durchlaufen die Indizes $\lambda, \mu, \nu, \omega$ die n Symbole I, II, ..., N, während $\alpha, \beta, \gamma, \delta$ außer I, II, ..., N auch noch das $(n+1)$-te Symbol 0 durchzulaufen haben.

[3]) Weyl, Zur Infinitesimalgeometrie: Einordnung der projektiven und der konformen Auffassung, Göttinger Nachrichten 1921, S. 99–112.

[4]) Schouten, Der Ricci-Kalkül, Berlin 1924, S. 76.

einem gegenüber (3) invarianten System Λ^k_{ij}:

$$\Lambda^k_{ij} = \Gamma^k_{ij} - \mathfrak{n}\,(A^k_i \Gamma^l_{lj} + A^k_j \Gamma^l_{li}) \qquad \left(\mathfrak{n} = \frac{1}{n+1}\right). \tag{4}$$

Die Koeffizienten Λ transformieren sich aber im allgemeinen nicht wie Γ und bestimmen deswegen keine Konnexion. Eine mit dem System Λ^k_{ij} versehene X_n soll mit $\mathfrak{A}_n$ bezeichnet werden.

Im folgenden werden wir zwei Probleme lösen:

I. Man soll in einer $\mathfrak{A}_n$ die Frenetschen Ableitungsformeln für eine gegebene Kurven*kongruenz* finden.

II. Dasselbe soll man in einer $\mathfrak{A}_n$ für eine gegebene *Kurve* machen.

Dabei verstehen wir unter „Frenetschen Formeln" ein System von Formeln, durch welches n mit der $\left\{\begin{matrix}\text{Kongruenz}\\ \text{Kurve}\end{matrix}\right.$ invariant verbundene Vektoren bestimmt werden, ganz analog wie in dem euklidischen Falle.

I.

1. Die Λ transformieren sich beim Koordinatenübergang $x^k \to x^\nu$ nach

$$\Lambda^k_{ij} = A^k_\nu A^\lambda_i A^\mu_j \Lambda^\nu_{\lambda\mu} + A^k_\lambda A^\mu_j \partial_\mu A^\lambda_i + \mathfrak{n}\,(A^k_j \partial_i \log A + A^k_i \partial_j \log A). \tag{5}$$

Dabei ist

$$A^k_i = A^k_\lambda A^\lambda_i, \qquad A = \text{Determinante von } A^k_\lambda, \quad \partial_i = \frac{\partial}{\partial x^i}.$$

Das letzte Glied rechts kommt von der Transformation

$$\Gamma^l_{li} = A^\mu_i (\Gamma^\lambda_{\lambda\mu} - \partial_\mu \log A) \tag{6}$$

her. Ist nun $\mathfrak{d}$ eine *beliebige* Dichte vom Gewicht -1 [5]), so folgt aus (5) und (6), daß sich die Koeffizienten

$$^*\Gamma^k_{ij} = \Lambda^k_{ij} + A^k_i \partial_j \log \mathfrak{d}^{-\mathfrak{n}} + A^k_j \partial_i \log \mathfrak{d}^{-\mathfrak{n}} \tag{7}$$

wie Γ transformieren, und bestimmen infolgedessen eine Konnexion. Da aber jede Dichte vom Gewicht -1 sich in der Form $p^{-\frac{1}{\mathfrak{n}}}\mathfrak{d}$ schreiben läßt, wenn p ein Skalar ist, so lehrt (7), daß die — durch $\mathfrak{A}_n$ induzierte — Konnexion nur *bis auf bahntreue Transformationen*

$$\overset{*}{\Gamma}{}^k_{ij} = {}^*\Gamma^k_{ij} + A^k_i p_j + A^k_j p_i \qquad (p_i = \partial_i \log p) \tag{3'}$$

bestimmt ist.

---

[5]) Eine Dichte vom Gewicht $\mathfrak{k}$ ist eine Größe mit einer einzigen Bestimmungszahl $\mathfrak{v}_{(0)}$, die bei dem Übergang $x^k \to x^\nu$ in $\mathfrak{v}_{(\Theta)}$ übergeht:

$$\mathfrak{v}_{(\Theta)} = A^{\mathfrak{k}}\,\mathfrak{v}_{(0)}.$$

Den Index (0), (Θ) lassen wir nachher immer fort.

2. Hier ist nun der Ort, wo wir uns der *gegebenen* Kongruenz bedienen werden, um uns von der *beliebigen* Dichte loszumachen. Den Tangentialvektor der Kongruenz bezeichnen wir mit v^k, und die zu den Konnexionen ${}^*\Gamma$ resp. $\overset{*}{\Gamma}$ gehörenden Symbole der kovarianten Derivation längs v^k mit ${}^*D$ resp. $\overset{*}{D}$. Dann konstruieren wir die Vektoren[6])

$$\overset{*}{v}_1 = v, \qquad \overset{*}{v}_w = \overset{*}{D}\,\overset{*}{v}_{w-1}, \tag{8}$$

$$ {}^*v_1 = v, \qquad {}^*v_w = {}^*D\,{}^*v_{w-1} \qquad (w = 2, \ldots, n). \tag{9}$$

Aus diesen Gleichungen beweist man leicht durch die Induktion von w auf $w+1$

$$\overset{*}{v}_w = {}^*v_w + r_{w-1} \qquad (w = 2, \ldots, n).$$

Dabei ist der Vektor r_{w-1} eine lineare Kombination von $v_1, \ldots, v_{w-1}$. Es folgt daraus für die beiden n-Vektoren ${}^*v^{i_1 \cdots i_n} = {}^*v_1^{[i_1} \ldots {}^*v_n^{i_n]}$ und $\overset{*}{v}{}^{[i_1 \cdots i_n]} = \overset{*}{v}_1^{[i_1} \ldots \overset{*}{v}_n^{i_n]}$

$${}^*v^{i_1 \cdots i_n} = \overset{*}{v}{}^{i_1 \cdots i_n}.$$

Der n-Vektor ${}^*v^{i_1 \cdots i_n}$ ist also eine Invariante gegenüber der bahntreuen Transformation (3′). Es ist also für die Konstruktion des n-Vektors ${}^*v^{i_1 \cdots i_n}$ belanglos, von welcher Dichte $\mathfrak{d}$ wir ausgehen. Mit anderen Worten, dieser n-Vektor ist eine Größe der $\mathfrak{A}_n$ und hängt von den — durch $\mathfrak{A}_n$ mittels (7) — induzierten Konnexionen nicht ab. Er kann auch identisch verschwinden, jedoch wollen wir voraussetzen, dies sei nicht der Fall, und unsere Kongruenz werden wir dann „allgemeine Kongruenz“ nennen. Die (bis auf das Vorzeichen) einzige, von Null verschiedene Bestimmungszahl dieses n-Vektors bezeichnen wir mit $\mathfrak{d}/n!$

$$\mathfrak{d} = {}^*v^{1 \cdots n}\, n! = \begin{vmatrix} {}^*v_1^1 & {}^*v_1^2 & \ldots & {}^*v_1^n \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ {}^*v_n^1 & {}^*v_n^2 & \ldots & {}^*v_n^n \end{vmatrix} = \overset{*}{v}{}^{1 \cdots n}\, n! = \begin{vmatrix} \overset{*}{v}{}_1^1 & \overset{*}{v}{}_1^2 & \ldots & \overset{*}{v}{}_1^n \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \overset{*}{v}{}_n^1 & \overset{*}{v}{}_n^2 & \ldots & \overset{*}{v}{}_n^n \end{vmatrix}.$$

$\mathfrak{d}$ ist eine Dichte vom Gewicht -1 und infolgedessen bestimmt sie, laut (7), eine Konnexion

$$\Gamma^k_{ij} = \Lambda^k_{ij} + A^k_i\, \partial_j \log \mathfrak{d}^{-n} + A^k_j\, \partial_i \log \mathfrak{d}^{-n}. \tag{10}$$

Die kovariante Ableitung von v in dieser Konnexion ist nach (10)

$$\partial_j \mathfrak{d} + \Gamma^k_{kj} \mathfrak{d} = 0.^{7}) \tag{11}$$

---

[6]) Die kontravarianten Indizes werden hier und überall, wo keine Verwirrung zu befürchten ist, unterdrückt.

[7]) Schouten-Hlavatý, Zur Theorie der allgemeinen linearen Übertragung, Math. Zeitschr. **30** (1929), S. 414–432.

Eine allgemeine Kongruenz in $\mathfrak{A}_n$ bestimmt eine Dichte vom Gewicht -1 und eine Konnexion, in welcher die kovariante Ableitung der Dichte identisch verschwindet[8]).

3. Die eben besprochene Konnexion enthält also kein *nicht* gegebenes Element und infolgedessen können wir uns ihrer bedienen, um die Frenetschen Formeln für die Kongruenz in $\mathfrak{A}_n$ abzuleiten. Zu dem Zwecke bezeichnen wir mit D das zur Konnexion (10) gehörende Symbol der kovarianten Derivation längs eines beliebig gegebenen Vektors e^k und führen folgende Vektoren ein:

$$\underset{1}{v} = \underset{1}{v}, \qquad \underset{w}{v} = D\,\underset{w-1}{v} \qquad (w = 2, \ldots, n), \tag{12a}$$

$$\underset{1}{e} = s\,\underset{1}{v}, \qquad \underset{w}{e} = D\,\underset{w-1}{e} \qquad (w = 2, \ldots, n). \tag{12b}$$

Hier ist s ein — später noch zu bestimmender — Skalar. Man überzeugt sich leicht, daß

$$\underset{[1}{e} \ldots \underset{n]}{e} = s^n \underset{[1}{v} \ldots \underset{n]}{v}. \tag{13}$$

Der n-Vektor $v^{i_1 \cdots i_n} = \underset{1}{v}^{[i_1} \ldots \underset{n}{v}^{i_n]}$ kann auch identisch verschwinden. Jedoch wollen wir zuerst $v^{i_1 \cdots i_n} \neq 0$ voraussetzen. Dann ist auch $e^{i_1 \cdots i_n} = \underset{1}{e}^{[i_1} \ldots \underset{n}{e}^{i_n]} \neq 0$. Die aus $v^{i_1 \cdots i_n}$ resp. $e^{i_1 \cdots i_n}$ entspringende Dichte vom Gewicht -1 wollen wir mit $\mathfrak{V}$ resp. $\mathfrak{e}$ bezeichnen,

$$\mathfrak{V} = n!\, v^{12 \ldots n}, \qquad \mathfrak{e} = n!\, e^{12 \ldots n},$$

und wählen den Skalar s nach der Vorschrift

$$s = \left(\frac{\mathfrak{v}}{\mathfrak{V}}\right)^{1/n}.$$

Dann ist laut (13)

$$\mathfrak{e} = \mathfrak{v}$$

und infolgedessen, nach (12b) und (11),

$$D\,\underset{[1}{e} \ldots \underset{n]}{e} = \underset{[1}{e} \ldots \underset{n-1}{e}\, D\,\underset{n]}{e} = 0.$$

Der Vektor $D\,\underset{n}{e}$ ist also eine lineare Kombination von $\underset{1}{e} \ldots \underset{n-1}{e}$.

$$D\,\underset{n}{e} = \alpha\,\underset{1}{e} + \ldots + \underset{n-1}{\alpha}\,\underset{n-1}{e}. \tag{14}$$

---

[8]) Mit jeder Wahl des Koordinatensystems ist gleichzeitig auch eine Dichte (vom Gewicht -1) gegeben, deren Bestimmungszahl gleich 1 ist. Dadurch ist laut (7) eine Konnexion ausgezeichnet, die in *diesem* Koordinatensystem die Koeffizienten

$$A^k_{ij} = \Gamma^k_{ij}$$

hat. Diese Konnexion hängt also von dem Koordinatensystem ab.

Die Skalaren α sind leicht zu berechnen. Bezeichnet man mit $\underset{u}{e}^{i_1 \dots i_n}$ den n-Vektor

$$\underset{u}{e}^{i_1 \dots i_n} = \underset{1}{e}^{[i_1} \dots \underset{u-1}{e}^{i_{u-1}} (D \underset{n}{e}^{i_u}) \underset{u+1}{e}^{i_{u+1}} \dots \underset{n}{e}^{i_n]}$$

und führt man die Dichte

$$\underset{u}{\mathfrak{e}} = \underset{u}{e}^{12 \dots n} n!$$

ein, so ist laut (14)

$$\boxed{\underset{u}{\alpha} = \underset{u}{\mathfrak{e}} / \mathfrak{v}}.$$

Die Formeln

$$\boxed{\begin{aligned} &D \underset{1}{e} = \underset{2}{e}, \quad D \underset{2}{e} = \underset{3}{e}, \quad \dots, \quad D \underset{n-1}{e} = \underset{n}{e}, \\ &D \underset{n}{e} = \underset{1\,1}{\alpha} \underset{1}{e} + \underset{2\,2}{\alpha} \underset{2}{e} + \dots + \underset{n-1\,n-1}{\alpha} \underset{n-1}{e} \end{aligned}}$$

kann man als Frenetsche Formeln einer allgemeinen Kongruenz in $\mathfrak{A}_n$ auffassen. Dann sind $\underset{1}{\alpha} \dots \underset{n-1}{\alpha}$ „Krümmungen" der Kongruenz in der beliebigen Richtung e^k.

4. Nun werden wir den früher angeschlossenen Fall behandeln, wo $v^{i_1 \dots i_n} = 0$. Da wir eine allgemeine Kongruenz betrachten, muß unbedingt in diesem Falle $e^k \neq v^k$. Der n-Vektor kann nur dann verschwinden, wenn sich unter den Vektoren $\underset{1}{v} \dots \underset{m}{v}\ \underset{m+1}{v} \dots \underset{n}{v}$ ein erster, sagen wir $\underset{m+1}{v}$ $(1 \leqq m < n)$ befindet, der eine lineare Kombination von den vorangehenden Vektoren $\underset{1}{v}, \dots, \underset{m}{v}$ ist. Dann existieren die Relationen

$$(15) \qquad D \underset{1}{v} = \underset{2}{v}, \quad \dots, \quad D \underset{m-1}{v} = \underset{m}{v}, \qquad D \underset{m}{v} = \underset{m+1}{v} = \underset{1\,1}{k} \underset{1}{v} + \dots + \underset{m\,m}{k} \underset{m}{v}$$

mit m Skalaren k. Ist $\underset{m}{k} = 0$, so kann man diese Formeln schon als Frenetsche Formeln mit $m - 1$ Krümmungen in der Richtung e^k auffassen. Wir setzen also voraus:

$$\underset{m}{k} \neq 0.$$

Um zu den Frenetschen Formeln zu gelangen, führen wir m Vektoren

$$\underset{1}{e} = s \underset{1}{v}, \quad \underset{2}{e} = D \underset{1}{e}, \quad \dots, \quad \underset{m}{e} = D \underset{m-1}{e}$$ [6])

ein. Dabei ist s ein — später noch zu bestimmender — Skalar. Es ist insbesondere

$$(16) \qquad \underset{u}{e} = s \underset{u}{v} + \underset{u-1}{r} \qquad (u = 1, \dots, m),$$

wenn $\underset{n-1}{r}$ eine lineare Kombination von den Vektoren $\underset{1}{v} \dots \underset{u-1}{v}$ ist und $\underset{0}{r} = 0$.

Anderseits ist nach (15)

$$D \underset{[1}{v} \dots \underset{m]}{v} = \underset{m\,[1}{k} \underset{}{v} \dots \underset{m]}{v},$$

und laut (16) und (15)

$$D\underset{[1}{e}\ldots\underset{m]}{e} = \underset{[1}{v}\ldots\underset{m]}{v}\,(e^k\partial_k \log s^m + \underset{m}{k})\,s^m$$

Nun denkt man sich e^k als Tangentialvektor eines Bahnkurvensystems einer eingliederigen Gruppe, so daß man schreiben kann

$$e^k = \frac{dx^k}{dt}.$$

Dann ist die Quadratur

$$Q = \int \underset{m}{k}\,dt$$

längs jeder allgemeinen Bahnkurve ausführbar und für

$$s = e^{-Q/m}$$

ist

$$D\underset{[1}{e}\ldots\underset{m]}{e} = 0 = \underset{[1\,2}{e}\,\underset{}{e}\ldots\underset{m-1}{e}\,D\underset{m]}{e}.$$

Der Vektor $D\underset{m}{e}$ ist also eine lineare Kombination von $\underset{1}{e},\ldots,\underset{m}{e}$,

$$D\underset{m}{e} = \alpha\underset{1\,1}{e} + \ldots + \underset{m-1\,m-1}{\alpha}\,\underset{}{e},$$

und die Skalaren sind mittels

$$\boxed{\underset{[1\,2}{e}\,\underset{}{e}\ldots\underset{u-1}{e}\,(D\underset{m}{e})\,\underset{u+1}{e}\ldots\underset{m]}{e} = \alpha\,\underset{u\,[1}{e}\ldots\underset{m]}{e}}$$

bestimmt. *Die Formeln*

$$\boxed{\begin{gathered}\underset{2}{e} = D\underset{1}{e},\quad \underset{3}{e} = D\underset{2}{e},\quad \ldots,\quad \underset{m}{e} = D\underset{m-1}{e},\\ D\underset{m}{e} = \alpha\underset{1\,1}{e} + \ldots + \underset{m-1\,m-1}{\alpha}\,\underset{}{e}\end{gathered}}$$

kann man in diesem Spezialfalle als Frenetsche Formeln einer allgemeinen Kongruenz in $\mathfrak{A}_n$ auffassen. Die $m-1$ Skalaren $\underset{1}{\alpha},\ldots,\underset{m-1}{\alpha}$ sind die Krümmungen der Kongruenz in der Richtung $\underset{1}{e} = v$.

Damit ist das erste Problem gelöst.

Bemerkung. Die aufgestellte Methode gilt nur im Falle einer allgemeinen Kongruenz. Ist die Kongruenz nicht allgemein, so kann sie im extremen Falle aus den in $\mathfrak{A}_n$ geodätischen Linien bestehen. Diese sind aber schon durch die Angabe von A bestimmt. In diesem Falle liefert also die Kongruenz kein neues Element, mittels welchem wir aus A eine ausgezeichnete Konnexion ableiten könnten. In allen anderen Fällen bestimmt die Kongruenz einen m-Vektor $(1 < m < n)$ der gegenüber $(3')$ invariant ist. Durch diesen m-Vektor ist eine nicht-holonome X_n^m gegeben

und man kann sie — mittels $\mathfrak{A}_n$ — zu einer $\mathfrak{A}_n^m$ ausstatten[9]). Nun liegt die Kongruenz in dieser $\mathfrak{A}_n^m$, wo sich also alles gerade so abspielt, wie früher in $\mathfrak{A}_n$. Man könnte sich also denken, das (jetzt nur m-dimensionale) Problem sei dadurch gelöst. Das ist aber keineswegs der Fall, weil die $\mathfrak{A}_n^m$ von der Einspannung in X_n abhängt. Das enthaltene Formelsystem hängt also noch von dieser *ganz beliebigen* Einspannung ab.

II.

1. Nun wollen wir das zweite Problem in Angriff nehmen, wo in $\mathfrak{A}_n$ nur *eine* Kurve C

$$x^k = x^k(t)$$

gegeben ist. Ganz analog wie früher kann man auch jetzt beweisen, daß mit der Kurve eine invariant verbundene Dichte $\mathfrak{v}$ existiert, welche von den bahntreuen Transformationen unabhängig ist. Im folgenden setzen wir stets $\mathfrak{v} \neq 0$ voraus und die Kurve nennen wir dann „allgemein“. Nun ist aber $\mathfrak{v}$ nur eine Funktion von t, so daß die Ableitungen von $\mathfrak{v}$ nach den Koordinaten sinnlos werden. Die Formel (10) ist jetzt unbrauchbar. Daher wollen wir einen anderen Weg einschlagen. Um den späteren Gedankengang nicht zu stören, führen wir an dieser Stelle einige neue, notwendige Begriffe ein.

Erstens führen wir längs C außer $x^1, \ldots, x^n$ noch eine überzählige Koordinate x^0 ein

$$x^0 = \log \mathfrak{v}^{-n/c} + \frac{t}{c} \tag{17}$$

mit einer beliebigen Konstante $c \neq 0$. *Zweitens* führen wir außer A_λ^k, A_j^ν noch die Koeffizienten

$$A_0^0 = 1, \qquad A_\lambda^0 = -\,n/c\,\partial_\lambda \log A, \qquad A_0^k = 0$$

ein. Es ist dann

$$|A_\beta^\alpha| = |A_\lambda^k| = A \neq 0\ ^{1}), ^{2}) \tag{18}$$

und infolgedessen lassen sich auch die inversen Koeffizienten A_α^α berechnen.

Die Differentiale dx^α transformieren sich also beim Übergang $x^k \to x^\nu$ nach

$$dx^\alpha = A_\alpha^\alpha\, dx^\alpha, \qquad dx^\alpha = A_\alpha^\alpha\, dx^\alpha.$$

Ein System von $n+1$ Bestimmungszahlen v^α (resp. w_α), welche sich kogredient (kontragredient) zu dx^α transformieren, soll ein kontravarianter (kovarianter) Punkt heißen. Multipliziert man einen solchen Punkt mit einer Dichte vom Gewicht $\mathfrak{k}$, so entsteht eine kontravariante (kovariante) Punktdichte $\mathfrak{v}^\alpha$ ($\mathfrak{w}_\alpha$) vom Gewicht $\mathfrak{k}$.

---

[9]) Schouten, Über nicht-holonome Übertragung in einer L_n, Math. Zeitschr. **30** (1929). In dieser Arbeit ist eine Methode zur Konstruktion von L_n^m angegeben. Analog kann auch eine $\mathfrak{A}_n^m$ in $\mathfrak{A}_n$ konstruiert werden.

Wir wollen jetzt die kontravarianten Punkte geometrisch veranschaulichen. Zu diesem Zwecke bemerken wir zuerst, daß die Bestimmungszahlen v^k eines solchen Punktes v^a sich nach

$$v^k = A_\lambda^k v^\lambda, \qquad v^\nu = A_k^\nu v^k$$

transformieren und deswegen einen kontravarianten Vektor v^k bestimmen. Dieser Vektor befindet sich in dem affin-ebenen Tangentialraume E_n von $\mathfrak{A}_n$. Wir denken uns also einen allgemeinen Punkt von $\mathfrak{A}_n$, bezeichnen ihn mit O und errichten im O den zugehörigen Raum E_n. Jeder Vektor p^k in $\mathfrak{A}_n$ findet seine Abbildung in dem Punkte von E_n, der die Koordinaten $y^k = p^k(0)$ hat und insbesondere entspricht dem Punkte O der feste Nullpunkt im E_n. Aus einem Punkte y^k kann man folgendermaßen einen kontravarianten Punkt v^a bilden: Man ordnet zuerst dem Punkte y^k einen skalaren Faktor $m(y) \neq 0$ zu und bildet nachher die Bestimmungszahlen v^k laut der Vorschrift

$$v^k = m\, y^k.$$

Um auch die letzte Bestimmungszahl v^0 definieren zu können, setzen wir *vorübergehend* die Ableitungen $\partial_k \mathfrak{v}$ als bekannt voraus und schreiben

$$v^0 = m/c\,(1 + y^k \cdot \partial_k \log \mathfrak{v}^{-n}) = m/c + v^k\, \partial_k \log \mathfrak{v}^{-n/c}.$$

(Man bestätigt leicht die richtige Transformationsweise von v^a.) Den Skalar m bezeichnen wir als „Masse" des kontravarianten Punktes mit den Bestimmungszahlen v^a. Ein kontravarianter Punkt ist also durch einen Punkt in E_n und durch die „Masse" vollkommen bestimmt. Umgekehrt läßt sich aus jedem kontravarianten Punkte mit der „Masse" $m \neq 0$ der zugehörige Punkt

$$v^k / (v^0 + v^i\, \partial_i \lg \mathfrak{v}^{n/c})\, c$$

in E_n konstruieren. Wir sagen, daß der kontravariante Punkt mit der Masse $m \neq 0$ sich auf einen Punkt in E_n abbildet.

Ist aber ein kontravarianter Punkt $v^a \neq 0$ gegeben, der *keine* „Masse" hat $(m = 0)$, so entspricht ihm in E_n kein Punkt. Vielmehr bestimmt er eine Richtung

$$v^1 : v^2 : \ldots : v^n,$$

wenn $v^k \neq 0$. — Dem kontravarianten Punkte $v^a = 0$, der auch keine „Masse" hat, entspricht in E_n keine Figur

Zur Abbildung von kontravarianten Punkten ist allgemein die Ableitung $\partial_k \log \mathfrak{v}$ nötig. Kennt man diese Ableitung nicht, so ist die Abbildungskonstruktion hinfällig. In diesem Falle bestimmt der kontravariante Punkt nur eine Richtung in E_n.

Wir kennen in unserem Falle die Ableitung $\partial_k \log \mathfrak{v}$ nicht und sind also nicht imstande für einen allgemeinen kontravarianten Punkt die geometrische Abbildung anzugeben. Die einzige Ausnahme bilden die kontra-

varianten Punkte, welche sich in die Tangentialrichtung von C abbilden. Ist O ein allgemeiner Kurvenpunkt, so gilt für jeden Punkt in E_n, der in die Tangentialrichtung fällt:

$$y^k = \varrho \underset{1}{e}{}^k, \qquad \underset{1}{e}{}^k = \frac{dx^k}{dt}.$$

Aus einem solchen Punkte läßt sich ein kontravarianter Punkt mit den Bestimmungszahlen

$$v^k = m\, y^k, \qquad v^0 = m/c\,(1 + y^k \partial_k \log \mathfrak{v}^{-\mathfrak{n}}) = m/c + m\,\varrho \frac{d}{dt} \log \mathfrak{v}^{-\mathfrak{n}/c}$$

konstruieren. Insbesondere läßt sich aus dem Kurvenpunkte O ein kontravarianter Punkt $\underset{0}{v}{}^a$ mit der „Masse“ 1 konstruieren. Für diesen Punkt ist $\varrho = 0$ zu setzen. Wir denken uns den Kurvenparameter t so gewählt, daß dem Spezialwerte $t = 0$ der Punkt O entspricht. Somit können wir für $\underset{0}{v}{}^a$ $\varrho = t$ setzen und erhalten

$$\underset{0}{v}{}^k = t\, \underset{1}{e}{}^k = 0, \qquad \underset{0}{v}{}^0 = 1/c + t \frac{d}{dt} \log \mathfrak{v}^{-\mathfrak{n}/c} = 1/c. \tag{19}$$

2. Ausgehend von Λ^i_{jk} sind die Herren Schouten und Gołąb zu einer projektiv-kovarianten Ableitung der Dichten $\mathfrak{v}^a$ (vom Gewicht $\mathfrak{n} + \mathfrak{k}$) resp. $\mathfrak{w}_a$ (vom Gewicht $-\mathfrak{n} - \mathfrak{k}$) gelangt,

$$\begin{aligned} D_b \mathfrak{v}^a &= \partial_b \mathfrak{v}^a + \overset{\mathfrak{k}}{\Lambda}{}^a_{cb}\, \mathfrak{v}^c, \\ D_b \mathfrak{w}_c &= \partial_b \mathfrak{w}_c - \overset{\mathfrak{k}}{\Lambda}{}^a_{cb}\, \mathfrak{w}_a, \end{aligned} \tag{20}$$

wo

$$\partial_0 = 0$$

und

$$\overset{\mathfrak{k}}{\Lambda}{}^i_{jk} = \Lambda^i_{jk}, \qquad \overset{\mathfrak{k}}{\Lambda}{}^c_{0j} = c\,\delta^c_j, \qquad \overset{\mathfrak{k}}{\Lambda}{}^c_{a0} = -c\,(n+1)\,\mathfrak{k}\,\delta^c_a, \text{ [10]}$$

$$\overset{\mathfrak{k}}{\Lambda}{}^0_{ij} = -1/c\,(n-1)\,(\partial_k \Lambda^k_{ij} - \partial_j \Lambda^k_{ik} + \Lambda^k_{lk}\Lambda^l_{ij} - \Lambda^k_{lj}\Lambda^l_{ik})$$

zu setzen ist. Führt man die Koeffizienten

$$\Pi^a_{cb} = \Pi^a_{bc} = 1/2 \left(\overset{-\mathfrak{n}}{\Lambda}{}^a_{bc} + \overset{-\mathfrak{n}}{\Lambda}{}^a_{cb} \right)$$

ein, so ergibt sich aus den obigen Gleichungen für beliebige Punkte v^a, w_a

$$D_b v^a = \partial_b v^a + \Pi^a_{cb}\, v^c, \qquad D_b w_a = \partial_b w_a - \Pi^c_{ab}\, w_c. \tag{21}$$

Da bei dem Koordinatenübergang $x^k \to x^\nu$ die Π sich nach

$$\Pi^c_{ab} = A^c_\gamma A^\alpha_a A^\beta_b \Pi^\gamma_{\alpha\beta} + A^c_\gamma \partial_b A^\gamma_a \tag{22}$$

[10]) Schouten-Gołąb, Über projektive Übertragungen und Ableitungen, Math. Zeitschr. 32 (1930). Ausgehend von den Koeffizienten A^α_a haben die Autoren die Punkte und Punktdichten durch die Transformationsweise definiert. Für diese Größen gibt es *allgemein kein* projektiv-kovariantes Differential. Durch die Einführung von x^0 läßt sich aber ein solches ausfindig machen, wie wir noch weiter beweisen.

transformieren, so ist D_a ein symbolischer kovarianter Punkt, und infolgedessen ist die Überschiebung

$$\Theta = dx^a D_a$$

ein symbolischer Skalar. Aus jedem kontravarianten Punkte v^a kann man also durch die projektiv-kovariante Derivation längs der Kurve einen neuen kontravarianten Punkt

$$\Theta v^a = dv^a + \Pi^a_{cb}\, dx^b\, v^c \tag{23}$$

konstruieren.

Die mit dem System $\overset{1}{\Lambda}{}^a_{bc}$ ausgestattete X_n soll $\mathfrak{P}_n$ heißen. Es gilt also der Satz: *Aus einer $\mathfrak{A}_n$ kann man eine $\mathfrak{P}_n$ konstruieren. Ist in dieser $\mathfrak{P}_n$ eine allgemeine Kurve gegeben, so kann man längs ihrer projektiv-kovariant differentiieren.*

3. Diesen Tatbestand wollen wir insofern ausnützen, indem wir aus dem gegebenen kontravarianten Kurvenpunkte $\underset{0}{v}$ [6]) n kontravariante Punkte $\underset{1}{v}, \ldots, \underset{n}{v}$ konstruieren

$$\underset{t}{\Theta}\, \underset{q}{v} = \underset{q+1}{v}, \qquad \underset{t}{\Theta} = \frac{dx^a}{dt} D_a \text{ [11])} \qquad (q = 0, \ldots, n-1). \tag{24}$$

Nun sollen diese Punkte $\underset{0}{v}, \ldots, \underset{n}{v}$ in bezug auf die Unabhängigkeit näher untersucht werden! Zu diesem Zwecke bemerken wir vorerst, daß jedem Punkte v^a der Vektor v^k in eindeutiger Weise zugeordnet ist. Insbesondere ist dem Punkte $\underset{t}{\Theta} v^a$ der Vektor $\underset{t}{\Theta} v^k$ zugeordnet. Bedient man sich der Gleichung (4), so kann man diesen Vektor in der Form

$$\underset{t}{\Theta} v^k = Dv^k + \omega\, \underset{1}{e}{}^k + \varepsilon v^k \tag{23'}$$

schreiben. Die darin auftretenden Skalaren ω, ε sind durch

$$\omega = c(v^0 - \mathfrak{n}/c\, \Gamma^i_{ij} v^j),$$
$$\varepsilon = c(\underset{1}{e}{}^0 - \mathfrak{n}/c\, \Gamma^i_{ij}\, \underset{1}{e}{}^j), \qquad \left(\underset{1}{e}{}^a = \frac{dx^a}{dt}\right)$$

---

[11]) Ist $m_{(q)}$ die Masse des Punktes $\underset{q}{v}$, so berechnet man die Masse $m_{(q+1)}$ des Punktes $\underset{q+1}{v}$ durch den Vergleich von

$$\underset{q+1}{v}{}^0 = \underset{t}{\Theta}\, \underset{q}{v}{}^0 = \frac{d}{dt}\, \underset{q}{v}{}^0 + \Pi^0_{ij}\, \underset{1}{e}{}^i\, \underset{q}{v}{}^j + c\, \underset{1}{e}{}^0\, \underset{q}{v}{}^0$$

und

$$\underset{q+1}{v}{}^0 = m_{(q+1)}/c + \underset{q+1}{v}{}^k\, \partial_k \log \mathfrak{v}^{-\mathfrak{n}/c}.$$

Es ist also

$$m_{(q+1)} = c\left(\frac{d}{dt}\, \underset{q}{v}{}^0 + \Pi^0_{ij}\, \underset{1}{e}{}^i\, \underset{q}{v}{}^j - \underset{q+1}{v}{}^k\, \partial_k \log \mathfrak{v}^{-\mathfrak{n}/c} + c\, \underset{1}{e}{}^0\, \underset{q}{v}{}^0\right).$$

Insbesondere ist für $t = 0$ und $q = 0$ nach dieser Gleichung und laut (24)

$$m_{(1)} = 1.$$

definiert, während der Vektor Dv^k folgendermaßen gegeben ist:

$$Dv^k = \underset{1}{e}^i(\partial_i v^k + \Gamma^k_{ji} v^j).$$

Wir haben also wieder eine durch $\mathfrak{A}_n$ induzierte Konnexion herausgegriffen, bemerken aber sogleich, daß dies *nur vorübergehend* geschehen ist, da wir mit Hilfe dieser Konnexion zu einem Resultat kommen, das von den bahntreuen Transformationen (3) unabhängig ist.

Insbesondere gibt uns (23′) für $v = \underset{0}{v}$ laut (19)

$$\underset{t\ 0}{\Theta} v^k = tD\underset{1}{e}^k + \underset{1}{e}^k + \underset{0}{\omega}\,\underset{1}{e}^k + t\varepsilon\,\underset{1}{e}^k, \tag{24′}$$

wenn

$$\underset{0}{\omega} = c(\underset{0}{v}^0 - \mathfrak{n}/c\,\Gamma^i_{ij}\underset{0}{v}^j).$$

Aus dem Vektor $\underset{1}{e}^k$ konstruieren wir die Vektoren

$$\underset{p}{e}^k = D\,\underset{p-1}{e}^k, \qquad (p = 2, \ldots, n+1), \tag{25}$$

so daß schließlich nach (24)

$$\underset{1}{v}^k = \underset{t\ 0}{\Theta} v^k = t\underset{2}{e}^k + (1 + \underset{0}{\omega} + t\varepsilon)\,\underset{1}{e}^k. \tag{26}$$

Wir setzen nun für ein ganzes q zwischen 1 und $n-1$ inklusive die Formel

$$\underset{q}{v}^k = t\,\underset{q+1}{e}^k + \underset{q}{e}^k(q + \underset{0}{\omega} + qt\varepsilon) + \underset{q-1}{U}^k \tag{27}$$

als bewiesen voraus. Dabei soll $\underset{q-1}{U}^k$ einen Vektor vorstellen, der sich in der $(q-1)$-Richtung von $\underset{1}{e}^k, \ldots, \underset{q-1}{e}^k$ befindet (und also $\underset{0}{U}^k = 0$).

Setzt man

$$\underset{q}{\omega} = c(\underset{q}{v}^0 - \mathfrak{n}/c\,\Gamma^i_{ij}\underset{q}{v}^j),$$
$$Q = (q + \underset{0}{\omega} + qt\varepsilon),$$

so folgt aus (23′) und (25) für $v^k = \underset{1}{v}^k$

$$\underset{q+1}{v}^k = t\,\underset{q+2}{e}^k + \underset{q+1}{e}^k[(q+1)(\varepsilon t + 1) + \underset{0}{\omega}]$$
$$+ \left\{\underset{c}{e}^k Q\left(\frac{d}{dt}\log Q + \varepsilon\right) + D\,\underset{q-1}{U}^k + \varepsilon\,\underset{q-1}{U}^k + \underset{1}{e}^k\,\underset{q}{\omega}\right\}.$$

Nach dem oben Gesagten ist der Vektor $D\,\underset{q-1}{U}^k$ in der q-Richtung von $\underset{1}{e}^k, \ldots, \underset{q}{e}^k$, und somit ist der in den geschweiften Klammern angeführte Vektor eine lineare Kombination von $\underset{1}{e}^k, \ldots, \underset{q}{e}^k$. Bezeichnet man diesen Vektor mit $\underset{q}{U}^k$, so kann man für die letzte Gleichung

$$\underset{q+1}{v}^k = t\,\underset{q+2}{e}^k + \underset{q+1}{e}^k[(q+1)(\varepsilon t + 1) + \underset{0}{\omega}] + \underset{q}{U}^k \tag{28}$$

schreiben. Der Vergleich von (26), (27) und (28) lehrt, daß (27) für alle Werte $q = 1, \ldots, n$ gültig ist. Diese Formel interessiert uns aber nur

in O. Dort ist $t = 0$ und außerdem nach (19)

$$\underset{0}{\omega} = 1. \tag{29}$$

Dadurch vereinfacht sich die Formel (27) zu

$$\underset{q}{v}^k = (q+1)\,\underset{q}{e}^k + \underset{q-1}{U}^k \qquad (q = 1, \ldots, n). \tag{30}$$

Es gilt somit in dem laufenden Punkte O

$$\underset{1}{v}^{[k_1} \ldots \underset{n}{v}^{k_n]} = \underset{1}{e}^{[k_1} \ldots \underset{n}{e}^{k_n]} (n+1)!\,.$$

Nun ist aber, wie wir schon früher (im ersten Abteil) zeigten, der n-Vektor $e^{k_1 \ldots k_n} = \underset{1}{e}^{[k_1} \ldots \underset{n}{e}^{k_n]}$ von der gebrauchten, beliebig eingeführten Konnexion *unabhängig*, und somit muß er für alle Konnexionen (3) entweder gleichzeitig identisch verschwinden, oder er kann für keine von diesen Konnexionen identisch verschwinden. Jedenfalls entspringt er aber aus dem *Tangential*vektor $\underset{1}{e}^k$, so daß er nur mit $\mathfrak{v}$ gleichzeitig verschwinden kann. Nach unserer Voraussetzung ist aber die Kurve C allgemein, d. i. $\mathfrak{v} \neq 0$, so daß auch der besprochene n-Vektor ungleich Null ist und infolgedessen, laut letzter Gleichung,

$$\mathfrak{v}^{k_1 \ldots k_n} = \underset{1}{v}^{[k_1} \ldots \underset{n}{v}^{k_n]} \neq 0.$$

Es gilt also für die Determinante

$$\mathfrak{w} = \begin{vmatrix} \underset{0}{v}^0 & \ldots & \underset{0}{v}^n \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \underset{n}{v}^0 & \ldots & \underset{n}{v}^n \end{vmatrix}, \tag{31}$$

laut (19), in dem laufenden Kurvenpunkte O

$$\mathfrak{w} = n!\,\underset{0}{v}^0\,\underset{0}{v}^{12 \ldots n} \neq 0$$

und somit *sind die durch* (24) *eingeführten Punkte samt* $\underset{0}{v}^a$ *linear unabhängig.*

Aus (31) und (18) folgt, daß $\mathfrak{w}$ eine Dichte vom Gewicht -1 ist.

4. Diese Dichte benützen wir zur Konstruktion eines ausgezeichneten Kurvenparameters. Zu dem Zwecke ordnen wir dem Punkte O eine andere „Masse“ zu, indem wir die Bestimmungszahlen $\underset{0}{v}^a$ mit einem — vorläufig ganz frei gewählten — Skalar s multiplizieren. Die neuen Bestimmungszahlen bezeichnen wir mit $\underset{0}{u}^a$

$$\underset{0}{u}^a = s\,\underset{0}{v}^a, \qquad s \neq 0. \tag{32}$$

Dann führen wir einen neuen Kurvenparameter $p(t)$ ein und definieren das Symbol

$$\underset{p}{\Theta} = \frac{dt}{dp}\underset{t}{\Theta} = \frac{dx^a}{dp} D_a.$$

Die aus $\underset{0}{u}{}^a$ entspringenden Punkte

$$\underset{i}{u}{}^a = \underset{p}{\Theta}\, \underset{i-1}{u}{}^a \qquad (i = 1, \ldots, n)$$

sind mit den Punkten $\underset{b}{v}{}^a$ durch

$$\underset{a}{u}{}^b = \left(\frac{dt}{dp}\right)^a s\, \underset{a}{v}{}^b + \underset{a-1}{V}{}^b$$

verbunden. Es gilt also

$$u^{a_0 \ldots a_n} = \underset{0}{u}{}^{[a_0} \ldots \underset{n}{u}{}^{a_n]} = \left(\frac{dt}{dp}\right)^{n(n+1)/2} s^{n+1}\, \underset{0}{v}{}^{[a_0} \ldots \underset{n}{v}{}^{a_n]}.$$

Die (bis auf das Vorzeichen) einzige Komponente von der ersten Größe links ist eine Dichte vom Gewicht -1 und wir wollen für sie das Symbol $\mathfrak{u}$ einführen:

$$\mathfrak{u} = n!\, u^{0 \ldots n}.$$

Die letzte Gleichung lehrt

$$\mathfrak{u} = \left(\frac{dt}{dp}\right)^{n(n+1)/2} s^{n+1}\, \mathfrak{w}.$$

Da $\mathfrak{w} \neq 0$, so ist auch $\mathfrak{u} \neq 0$ und die Punkte $\underset{b}{u}{}^a$ sind linear unabhängig. Nun wählen wir

$$(33) \qquad p = \int \left(\frac{\mathfrak{v}\, e^{-(n+1)t}}{\mathfrak{w}\, s^{n+1}}\right)^{-2/n(n+1)} dt = \int \left[\left(\frac{\mathfrak{v}}{\mathfrak{w}}\right)^n \frac{1}{s e^t}\right]^{-2/n} dt,$$

so daß

$$(34\,a) \qquad \mathfrak{u} = \mathfrak{v}\, e^{-(n+1)t}.$$

Die Gleichung (23) lehrt

$$\underset{t}{\Theta} u^{a_0 \ldots a_n} = \frac{d}{dt} u^{a_0 \ldots a_n} + \Pi^a_{ab}\, \underset{1}{e}{}^b\, u^{a_0 \ldots a_n} = \frac{d}{dt} u^{a_0 \ldots a_n} + \frac{c}{\mathfrak{n}}\left(\frac{1}{c} - \frac{\mathfrak{n}}{c}\frac{d \log \mathfrak{v}}{dt}\right) u^{a_0 \ldots a_n},$$

so daß laut (34a)

$$(34\,b) \qquad \underset{p}{\Theta}\, u^{a_0 \ldots a_n} = \underset{0}{u}{}^{[a_0} \ldots \underset{n-1}{u}{}^{a_{n-1}}\, \underset{p}{\Theta}\, \underset{n}{u}{}^{a]} = 0,$$

und der Punkt $\underset{p}{\Theta}\, \underset{n}{u}{}^a$ eine lineare Kombination von $\underset{0}{u}{}^a \ldots \underset{n-1}{u}{}^a$ ist:

$$(35) \qquad \underset{p}{\Theta}\, \underset{n}{u}{}^a = \sum_0^{n-1} \underset{w}{\overset{w}{\varkappa}}\, \underset{w}{u}{}^a \qquad (w = 0, \ldots, n-1).$$

Um die Skalare $\varkappa$ zu berechnen, führen wir einfachheitshalber noch die Punkte

$$\underset{n+q}{u}{}^a = \underset{p}{\Theta}\, \underset{n+q-1}{u}{}^a$$

ein. Dann ergibt sich aus (35)

$$(36) \qquad (-1)^n\, \underset{[1}{u} \ldots \underset{n+1]}{u} = \varkappa\, \underset{[0}{u} \ldots \underset{n]}{u},$$

$$\underset{[0}{u} \ldots \underset{r-1}{u}\, \underset{n+1}{u}\, \underset{r+1}{u} \ldots \underset{n]}{u} = \overset{r}{\varkappa}\, \underset{[0}{u} \ldots \underset{n]}{u} \qquad (r = 1, \ldots, n-1).$$

Die kovariant-projektive Derivation lehrt nach (34 b)

$$(37)\quad \overset{r-1}{-\varkappa}\, \underset{[0}{u} \dots \underset{n]}{u} = \underset{[0}{u} \dots \underset{n]}{u} \frac{d}{dp} \overset{r}{\varkappa} + (-1)^{n-r+1} \underset{[0}{u} \dots \underset{r-1}{u}\, \underset{r+1}{u} \dots \underset{n}{u}\, \underset{n+2]}{u} \qquad (r>1).$$

Mindestens eine von den Gleichungen (36), (37) läßt sich durch spezielle Wahl von dem noch frei gebliebenen Skalar s vereinfachen

5. Wenn mindestens eine von den rechts angeschriebenen Größen ungleich Null ist, z. B

$$\underset{[0}{u} \dots \underset{q-1}{u}\, \underset{q+1}{u} \dots \underset{n}{u}\, \underset{n+2]}{u} \neq 0 \qquad (q>1),$$

so bezeichnen wir die daraus entspringende Dichte vom Gewicht -1 mit $\mathfrak{q}$

$$\mathfrak{q} = \underset{[0}{u^0} \dots \underset{q-1}{u^{q-1}}\, \underset{q+1}{u^q} \dots \underset{n}{u^{n-1}}\, \underset{n-2]}{u^n}\, n!$$

so daß

$$(38)\quad \overset{q-1}{-\varkappa} = \frac{d}{dp} \overset{q}{\varkappa} + (-1)^{n-q+1} \mathfrak{q}/\mathfrak{u}.$$

Nun ordnen wir dem Kurvenpunkte O eine andere „Masse“ zu, indem wir statt s einen neuen Faktor s' einführen. Für alle zu diesem Faktor gehörenden Elemente gebrauchen wir dieselbe Bezeichnung (wie für s), nur mit Akzent versehen, und das Verhältnis s'/s bezeichnen wir mit σ. Es folgt dann aus

$$\underset{0}{u} = s \underset{0}{v}, \qquad \underset{0}{u'} = s' \underset{0}{v}$$

die Relation

$$\underset{0}{u'} = \sigma \underset{0}{u}.$$

Anderseits bekommt man aus (33) und aus der analogen Gleichung für s'

$$\frac{dp}{dp'} = \sigma^{-2/n}.$$

Die Punkte

$$\underset{0}{u'},\ \underset{i}{u'} = \underset{p}{\Theta}\, \underset{i-1}{u'} = \frac{dp}{dp'} \underset{p}{\Theta}\, \underset{i-1}{u'} \qquad (i=1, \dots, n)$$

sind mit $\underset{a}{u}$ mittels

$$(39)\quad \underset{a}{u'} = \sigma^{(-2a+n)/n} \underset{a}{u} + \underset{a-1}{U}, \qquad \underset{0}{U} = 0$$

verbunden. Dabei ist $\underset{a-1}{U}$ eine lineare Kombination von $\underset{0}{u}, \dots, \underset{a-1}{u}$, und die darin auftretenden Koeffizienten sind Funktionen nur von σ und von den Ableitungen von σ nach p. Die Punkte $\underset{n}{u'}$ sind also auch linear unabhängig. Aus (39) folgt, daß man die Dichte $\mathfrak{q}'$ als eine Summe von Dichten (alle von demselben Gewichte -1) schreiben kann, und daß die darin auftretenden Koeffizienten *nur* von σ und von den Ableitungen von σ nach p abhängen. Diese Ableitungen kann man nach (33) in die Ableitungen nach t umsetzen. Die kovariante Gleichung

$$(40)\quad \mathfrak{q}' = (-1)^{n-q+1} \mathfrak{v}\, e^{-t/n}$$

ist also eine Differentialgleichung für σ mit der Lösung

$$\sigma = s'/s = f(t, s).$$

Setzt man nun überall $s = 1$ und läßt man die Akzente nachher wieder weg, so ergibt sich unter der Voraussetzung $f(t, 1) \neq 0$ für den ausgezeichneten „projektiven Kurvenbogen“

$$p = \int \left(\frac{e^{-(n+1)t}\, \mathfrak{v}(t)}{\mathfrak{w}(t)\, f(t,1)^{n+1}} \right)^{-2/n(n+1)} dt = \int \left[\left(\frac{\mathfrak{v}(t)}{\mathfrak{w}(t)} \right)^{n} \frac{1}{e^{t} f(t,1)} \right]^{-2/n} dt$$

laut (38) und (40)

$$-\overset{q-1}{\varkappa} = \frac{d}{dt} \overset{q}{\varkappa} + 1$$

und es gilt also

$$\begin{gathered} \underset{p}{\Theta}\, \underset{0}{u} = \underset{1}{u}, \ldots, \underset{p}{\Theta}\, \underset{n-1}{u} = \underset{n}{u}, \\ \underset{p}{\Theta}\, \underset{n}{u} = \overset{0}{\varkappa}\, \underset{0}{u} + \ldots + \overset{q-2}{\varkappa}\, \underset{q-2}{u} - \left(\frac{d}{dt} \overset{q}{\varkappa} + 1 \right) \underset{q-1}{u} + \overset{q}{\varkappa}\, \underset{q}{u} + \ldots + \overset{n-1}{\varkappa}\, \underset{n-1}{u}. \end{gathered}$$

Diese Gleichungen mit $n-1$ Skalaren $\varkappa$ kann man als Frenetsche Formeln einer allgemeinen Kurve in $\mathfrak{P}_n$ auffassen. Die Skalaren $\varkappa$ sind dann projektive „Krümmungen“.

6. Wenn alle in (37) rechts auftretenden Größen Null sind, so läßt sich die Formel

$$\frac{d^r}{dp^r} \overset{r+1}{\varkappa} = (-1)^r \overset{1}{\varkappa} \qquad (r = 1, \ldots, n-2)$$

leicht beweisen. Dabei wird aber p nur durch (33) definiert. Ist außerdem $\underset{\{1}{u} \ldots \underset{n+1\}}{u} \neq 0$, so läßt sich wieder ein ausgezeichneter projektiver Kurvenbogen p angeben, für welchen $\overset{0}{\varkappa} = 1$.

7. Neben den Frenetschen Formeln für kontravariante Punkte lassen sich auch „duale“ Formeln für kovariante Punkte angeben. Zu diesem Zwecke konstruiert man zuerst nach

$$\overset{a}{u}_c \underset{b}{u}^c = \delta^a_b$$ [12])

---

[12]) Die kovarianten Punkte lassen sich auf E_n folgendermaßen abbilden:

$$w_k = y_k + M\, \partial_k \log \mathfrak{v}^{n/e}, \qquad w_0 = M.$$

Dabei sind y_k die Bestimmungszahlen einer Hyperebene in E_n (eines kovarianten Vektors), und das Skalarfeld $M(y)$ kann als „Masse“ des kontravarianten Punktes w_a bezeichnet werden. Insbesondere ist also ein kovarianter Punkt ohne Masse mit einer Hyperebene in E_n identisch. — Zwei kovariante Punkte w_a und ϱw_a haben verschiedene Abbildungshyperebenen (mit der gemeinsamen $(n-1)$-Richtung), während zwei kontravariante Punkte v^a, ϱv^a denselben Abbildungspunkt haben. Zur Abbildung eines kovarianten Punktes ist also die Ableitung $\partial_k \mathfrak{v}$ nötig.

die Punkte $\overset{a}{u}_c$. Aus dieser Gleichung folgt

$$\left(\underset{p}{\Theta}\overset{a}{u}_c\right)\overset{c}{u}_b + \overset{a}{u}_c \underset{p}{\Theta}\overset{c}{u}_b = 0$$

und infolgedessen

$$\boxed{\underset{p}{\Theta}\overset{a}{u}_c = -\left(\overset{a-1}{u}_c + \overset{n}{u}_c \overset{a}{\varkappa}\right)} \qquad \left(\overset{-1}{u}_c = 0,\ \overset{n}{\varkappa} = 0\right).$$

Das sind die gesuchten dualen Frenetschen Formeln für kontravariante Punkte längs C. Kennt man die Krümmungen $\varkappa$, so lassen sich aus $\overset{0}{u}_c$ die übrigen Punkte $\overset{i}{u}_a$ mittels dieser Formeln ableiten.

8. Jedem kovarianten Punkte w_a kann man die Skalare

$$\underset{a}{w} = w_b \overset{b}{u}_a$$

zuordnen. Es ist dann laut den oben aufgestellten Formeln

$$\underset{p}{\Theta}\, w_b = \left(\underset{p}{\vartheta}\,\underset{a}{w}\right)\overset{a}{u}_b,$$

wo

$$\underset{p}{\vartheta}\,\underset{b}{w} = \frac{d}{dp}\,\underset{b}{w} - \delta_b^{d-1}\,\underset{d}{w} - \delta_b^{n}\,\underset{d}{w}\,\overset{d}{\varkappa} \qquad \left(\overset{n}{\varkappa} = 0\right). \tag{41}$$

Die kovariante Ableitung $\underset{p}{\Theta}\, w_a$ *von* w_a *läßt sich immer auf die einfachere Form* (41) *bringen.*

Analogisch kann man jedem kontravarianten Punkte v^a die Skalare

$$\overset{a}{v} = v^b \underset{b}{u}^a$$

zuordnen, so daß

$$v^a = \overset{b}{v}\,\underset{b}{u}^a.$$

Es gilt dann der Satz: *Die kovariante Ableitung* $\underset{p}{\Theta}\, v^a$ *von* v^a *läßt sich immer auf die einfachere Form*

$$\underset{p}{\vartheta}\,\overset{b}{v} = \frac{d}{dp}\,\overset{b}{v} + \overset{b-1}{v} + \overset{n}{v}\,\overset{b}{\varkappa} \qquad \left(\overset{-1}{v} = \overset{n}{\varkappa} = 0\right) \tag{42}$$

bringen.

In beiden Fällen ist natürlich p als ausgezeichneter projektiver Kurvenbogen gedacht und für die $\varkappa$ ist eine von den oben erwähnten Möglichkeiten einzusetzen.

(Eingegangen am 24. Oktober 1930.)

Über die Abplattung der homogenen Gleichgewichtsfiguren rotierender, gravitierender Flüssigkeiten.

Von

Władysław Nikliborc in Lwów, z. Z. in Leipzig[1]).

Es sei T eine Gleichgewichtsfigur homogener, gravitierender Flüssigkeit, welche mit der konstanten Winkelgeschwindigkeit ω um die z-Achse rotiert. Der Schwerpunkt von T liegt dann auf der z-Achse, und die in diesem Punkte auf der z-Achse senkrecht stehende Ebene ist zugleich eine Symmetrieebene von T[2]). Es sei jetzt $P(x_0, y_0, b)$ derjenige Punkt von T, in welchem die z-Koordinate ihren größten Wert annimmt. Es bedeute weiter a den größten Wert von $\sqrt{x^2+y^2}$ in T. Die Größe

$$\frac{b}{a}$$

wollen wir als „*Abplattung*“ von T bezeichnen.

Man verdankt dem Herrn Mazurkiewicz[3]) eine Abschätzung dieser Größe. Er bewies, daß entweder

$$b < a$$

oder

$$b < e^{168050}\, v^{1/3}$$

ist, wobei v das Volumen von T bedeutet. Bedenkt man, daß

$$v < 2a^2 b\pi$$

[1]) International Education Board Fellow (Rockefeller Foundation).

[2]) L. Lichtenstein, Über einige Eigenschaften der Gleichgewichtsfiguren rotierender, homogener Flüssigkeiten, deren Teilchen einander nach dem Newtonschen Gesetz anziehen, Sitzungsber. der Preuß. Akad. der Wiss. 1918, Erster Hlbbd., S. 1120—1135, insbes. S. 1123.

[3]) Vgl. W. Mazurkiewicz, Zur Theorie der Gleichgewichtsfiguren rotierender, homogener Flüssigkeiten, Math. Zeitschr. 25 (1926), S. 749—753.

ist, so folgt hieraus sofort eine Abschätzung der Abplattung, die allerdings sehr roh ist[4]) [5]).

In dieser Arbeit beschäftigen wir uns mit der Abschätzung der Größe $\frac{b}{a}$ *in dem Spezialfalle einer Rotationsfigur vom Kugeltypus, welche darüber hinaus noch die Eigenschaft hat, daß die Meridiankurve im Halbraume* $z > 0$ *monoton verläuft. In diesem Falle wird gezeigt, daß*

$$\frac{b}{a} < 10$$

gilt.

§ 1.

Reihenentwicklung des Volumenpotentials.

1. In diesem Paragraphen wird es sich um eine Reihenentwicklung des von einem Rotationskörper T herrührenden Newtonschen Volumenpotentials handeln. Die Dichte wird dabei als konstant (und $= 1$) vorausgesetzt. Des weiteren setzen wir folgendes voraus:

Jede zur Rotationsachse parallele oder senkrechte Gerade soll den Rand S von T höchstens in zwei Punkten schneiden, oder mit demselben eine ganze Strecke gemeinsam haben.

Hieraus folgt natürlich sofort, daß T vom topologischen Kugeltypus sein muß.

Wählen wir die z-Achse zur Rotationsachse, so folgt weiter aus unserer Voraussetzung, daß sich die Gleichung des Mantels von T auf die Form

$$x^2 + y^2 = \varrho^2(z) \qquad (\varrho(z) \geqq 0)$$

bringen läßt, wobei $\varrho(z)$ den folgenden Bedingungen genügen muß:

1. In dem Intervall (b_1, b_2), in dem $\varrho(z)$ erklärt ist, ist diese Funktion eindeutig und stetig.

2. $\varrho(z)$ besitzt im Innern von (b_1, b_2) kein Minimum (es sei denn, daß überhaupt $\varrho(z) = \text{konst.}$ ist).

Es seien nun $M(x, y, z)$ und $P(\xi, \eta, \zeta)$ zwei Punkte, r_{PM} ihr Abstand. Wir betrachten das Newtonsche Volumenpotential

$$V(P) = \int_T \frac{d\tau_P}{r_{PM}}.$$

---

[4]) Es sei noch bemerkt, daß Herr E. Hölder im Anschluß an die Arbeit des Herrn Mazurkiewicz bewiesen hat, daß entweder $b \leqq a$ oder $b < e^{9000}\, v^{1/3}$ gilt. Hieraus folgt eine bessere Abschätzung für $\frac{b}{a}$, die jedoch noch immer sehr roh ist. Vgl. E. Hölder, Math. Zeitschr. **25** (1926), S. 754.

[5]) Die Größe a wurde von Herrn Lichtenstein (loc. cit. [2])) abgeschätzt.

Die Integration ist über ganz T zu erstrecken; $d\tau_P$ bedeutet das Volumenelement im Punkte P. Bezeichnen wir mit $K(z)$ die Kreisfläche

$$x^2+y^2 \leqq \varrho^2(z), \qquad z=0,$$

so ist

$$V(P)=\int_{b_1}^{b_2}\left\{\iint_{K(z)}\frac{dx\,dy}{\sqrt{(x-\xi)^2+(y-\eta)^2+(z-\zeta)^2}}\right\}dz.$$

Um dieses Integral in eine unendliche Reihe zu entwickeln, führen wir die ebenen Polarkoordinaten (t, φ)

$$x=t\cos\varphi, \qquad y=t\sin\varphi$$

ein. Ohne etwas an der Allgemeinheit zu verlieren, können wir natürlich $\eta=0$ setzen. Es ist dann

$$(1)\qquad V(P)=\int_{b_1}^{b_2}dz\int_0^{\varrho(z)}dt\int_0^{2\pi}\frac{t\,d\varphi}{\sqrt{t^2-2\xi t\cos\varphi+\xi^2+(\zeta-z)^2}}.$$

2. Wir wollen jetzt die für

$$|u|<1$$

gültige Formel

$$\int_0^{2\pi}\frac{d\varphi}{\sqrt{1+u\cos\varphi}}=2\pi+2\pi\sum_{k=1}\frac{1\cdot3\cdot\ldots\cdot4k-1}{(2k)!}\cdot\frac{1\cdot3\cdot\ldots\cdot2k-1}{2\cdot4\cdot\ldots\cdot2k}\cdot\left(\frac{u}{2}\right)^{2k},$$

zu der man durch eine einfache Anwendung der Binomialreihe für $\frac{1}{\sqrt{1+x}}$ und gliedweise Integration gelangt, benutzen.

Betrachten wir die Funktion $A(\varrho)$

$$(2)\qquad A(\varrho)=\int_0^{\varrho}\left\{\int_0^{2\pi}\frac{d\varphi}{\sqrt{t^2-2\xi t\cos\varphi+\xi^2+(\zeta-z)^2}}\right\}t\,dt.$$

Wegen $A(0)=0$ ist offenbar

$$A(\varrho)=\int_0^{\varrho}A'(\varrho)\,d\varrho,$$

$$A'(\varrho)=\int_0^{2\pi}\frac{\varrho\,d\varphi}{\sqrt{\varrho^2-2\xi\varrho\cos\varphi+\xi^2+(\zeta-z)^2}}.$$

Ist nun $\zeta+z$, so ist gewiß

$$\frac{2|\xi\varrho|}{\varrho^2+\xi^2+(\zeta-z)^2}<1,$$

und daher

$$(3) \qquad A(\varrho) = 2\pi \int_0^{\varrho} \frac{t\,dt}{\sqrt{t^2+\xi^2+(\zeta-z)^2}}$$

$$+ 2\pi \sum_{k=1}^{\infty} \frac{1\cdot 3\cdot \ldots \cdot (4k-1)}{(2k)!} \cdot \frac{1\cdot 3\cdot \ldots \cdot (2k-1)}{2\cdot 4\cdot \ldots \cdot 2k} \cdot \int_0^{\varrho} \frac{\xi^{2k} t^{2k+1} dt}{[t^2+\xi^2+(\xi-z)^2]^{2k+\frac{1}{2}}}.$$

Wir gehen jetzt zu dem Falle

$$z = \zeta$$

über. Zunächst die Definition (2) von $A(\varrho)$ selbst. Das rechter Hand stehende Integral ist gewiß konvergent, da es als Potential der einfachen Belegung, welches von der Kreisfläche

$$x^2 + y^2 = \varrho^2, \qquad z = \zeta$$

herrührt, betrachtet werden kann. Dieses ist aber eine im ganzen Raume stetige Funktion.

Um jetzt die Gültigkeit von (3) auch für $z = \zeta$ zu beweisen, bemerken wir zunächst, daß das Integral

$$J(\xi, \zeta, z) = \int_0^{\varrho} \frac{\xi^{2k} t^{2k+1} dt}{[t^2+\xi^2+(\zeta-z)^2]^{2k+\frac{1}{2}}}$$

von z, ζ und ξ stetig abhängt. Dies ist, wie man es durch Integration leicht bestätigt, für $k = 0$ ohne weiteres ersichtlich. Für $k > 0$ ist dieses Integral stetig für alle Wertetrippel (ξ, ζ, z), die der Ungleichheit

$$|\xi| + |\zeta - z| \neq 0$$

genügen. Es bleibt uns also noch die Stetigkeit des Integrals für diejenigen Wertetrippel (ξ, ζ, z) nachzuweisen, für welche

$$\xi = 0, \qquad \zeta = z$$

gilt.

Wegen $J(0, \zeta, z) = 0$ ist für $\zeta = z$

$$J(\xi, z+h, z+l) - J(0, z, z) = \int_0^{\varrho} \frac{\xi^{2k} t^{2k+1} dt}{[t^2+\xi^2+(h-l)^2]^{2k+\frac{1}{2}}},$$

also nach der Substitution $t = \xi u$

$$|J(\xi, z+h, z+l) - J(0, z, z)| = \left| \int_0^{\frac{\varrho}{\xi}} \frac{\xi^{4k+2} u^{2k+1} du}{[\xi^2(1+u^2)+(h-l)^2]^{2k+\frac{1}{2}}} \right|$$

$$\leqq \left| \int_0^{\frac{\varrho}{\xi}} \frac{\xi^{4k+2} u^{2k+1} du}{[\xi^2(1+u^2)]^{2k+\frac{1}{2}}} \right| \leqq \left| \xi \cdot \int_0^{\frac{\varrho}{\xi}} \frac{u^{2k+1} du}{(1+u^2)^{2k+\frac{1}{2}}} \right| < |\xi| \cdot \int_0^{\infty} \frac{u^{2k+1} du}{(1+u^2)^{2k+\frac{1}{2}}} = |\xi|\, C,$$

wo C eine Konstante bedeutet. Hieraus folgt offenbar die Stetigkeit der Funktion $J(\xi, \zeta, z)$ für $\xi = 0$, $\zeta = z$.

Es ist jetzt klar, daß, sobald die gleichmäßige Konvergenz der in (3) vorkommenden Reihe festgestellt ist, ist zugleich gesichert, daß die Summe eine stetige Funktion von ξ und ζ darstellt. Weiter aber folgt dann, daß die Formel (3) auch für

$$z = \zeta$$

gilt, da sowohl die linke als auch die rechte Seite stetig ist, und die Formel für $z \neq \zeta$ bereits bewiesen wurde.

Wir sehen also: um die Gültigkeit von (3) allgemein zu beweisen, genügt es, die gleichmäßige Konvergenz der dort vorkommenden Reihe festzustellen.

3. Dies geschieht folgenderweise: Wir setzen zunächst $\zeta = z$, wodurch sämtliche Glieder von (3) vergrößert werden, und führen dann eine neue Variable τ durch die Substitution

$$t = \xi \frac{\sqrt{1-\tau^2}}{\tau}$$

ein. Man bekommt auf diese Weise die Reihe

$$2\pi \sum_{k=1}^{\infty} \frac{1\cdot 3\cdot \ldots \cdot 4k-1}{(2k)!} \cdot \frac{1\cdot 3\cdot \ldots \cdot 2k-1}{2\cdot 4\cdot \ldots \cdot 2k} \cdot |\xi| \cdot \int_{\tau^*}^{1} \tau^{2k-2}(1-\tau^2)^k d\tau,$$

wo $0 < \tau^* < 1$ ist. Nun ist

$$\int_{\tau^*}^{1} \tau^{2k-2}(1-\tau^2)^k d\tau \leqq \int_{0}^{1} \tau^{2k-2}(1-\tau^2)^k d\tau.$$

Das Integral rechter Hand berechnet man sofort durch sukzessive partielle Integration mit $u = (1-\tau^2)^k$, $dv = \tau^{2k-2} d\tau$, und erhält den Wert

$$\frac{2k}{2k-1} \cdot \frac{2k-2}{2k+1} \cdot \ldots \cdot \frac{2}{4k-3} \cdot \frac{1}{4k-1}.$$

Somit ist die Reihe

$$(4) \quad |\xi| \cdot \sum_{k=1}^{\infty} \frac{1\cdot 3\cdot \ldots \cdot 4k-1}{(2k)!} \cdot \frac{1\cdot 3\cdot \ldots \cdot (2k-1)}{2\cdot 4\cdot \ldots \cdot 2k} \cdot \frac{2\cdot 4\cdot \ldots \cdot 2k}{(2k-1)\cdot(2k+1)\cdot \ldots \cdot (4k-3)} \cdot \frac{1}{4k-1}$$

eine Majorante für (3). Von der Konvergenz der Reihe (4) kann man sich aber sofort mit Hilfe des Raabeschen Kriteriums überzeugen.

Damit ist die gleichmäßige Konvergenz von (3) in jedem beschränkten Bereiche bewiesen.

4. Unter Benutzung der Formel (3) bekommen wir aus (1) sofort die gesuchte Reihenentwicklung für $V(P)$.

$$(5) \qquad V(P) = 2\pi \int_{b_1}^{b_2} dz \int_0^{\varrho(z)} \frac{t\,dt}{\sqrt{t^2+\xi^2+(\zeta-z)^2}}$$

$$+ 2\pi \sum_{k=1}^{\infty} \frac{1\cdot 3\cdot \ldots \cdot 4k-1}{(2k)!} \cdot \frac{1\cdot 3\cdot \ldots \cdot (2k-1)}{2\cdot 4\cdot \ldots \cdot 2k} \cdot \int_{b_1}^{b_2} dz \int_0^{\varrho(z)} \frac{\xi^{2k} t^{2k+1}\,dt}{[t^2+\xi^2+(\zeta-z)^2]^{2k+\frac{1}{2}}}.$$

§ 2.

Problemstellung.

5. Wir beginnen damit, daß wir den in § 1 aufgezählten Eigenschaften von T noch die folgende neue Voraussetzung hinzufügen.

T soll eine auf der z-Achse senkrecht stehende Symmetrieebene besitzen.

Wird als diese die $z = 0$-Ebene gewählt, so folgt daraus, daß die Funktion $\varrho(z)$ die folgenden zwei Eigenschaften hat:

1. Es ist
$$\varrho(z) = \varrho(-z).$$

2. Setzt man zur Abkürzung
$$b_1 = b, \qquad b_2 = -b,$$
so ist die Funktion $\varrho(z)$ in dem Intervall $(0, b)$ abnehmend.

Wir bezeichnen mit a irgendeine positive Zahl und betrachten die Menge $\mathfrak{M}$ aller Figuren T, welche

1. alle bisher aufgezählten Eigenschaften haben und
2. in dem Bereiche
$$x^2 + y^2 \leqq a^2,$$
$$|z| \leqq b$$
enthalten sind.

Es sei jetzt noch ζ irgendeine der Ungleichheit
$$0 < \zeta < b$$
genügende Zahl.

Wir betrachten irgendeine zu $\mathfrak{M}$ gehörende Figur T und die beiden Punkte P, Q mit den Koordinaten $\{0, 0, b\}$ und $\{\varrho(\zeta), 0, \zeta\}$. Es sei $V_T(B)$ der Wert des von T herrührenden Potentials im Punkte B. Wir betrachten die Differenz der Potentiale
$$V_T(P) - V_T(Q)$$
und stellen uns die Aufgabe:

Es soll T_0 in $\mathfrak{M}$ so bestimmt werden, daß diese Differenz das (absolute) Maximum wird.

Es handelt sich hier um ein eigenartiges Problem der Variationsrechnung. Es ist nicht allein eine zur Menge $\mathfrak{M}$ gehörende Figur, oder wenn man es will, eine Fläche S, oder aber eine Funktion $\varrho(z)$, unbekannt, sondern es wird darüber hinaus ein Punkt auf dieser Fläche, der auf dem Kreise

$$x^2 + y^2 = a^2,$$
$$z = \zeta$$

liegen soll, gesucht.

Es liegt hier ein Maximum eines Integrals vor, welches nicht nur von der unbekannten Funktion abhängt, sondern dessen Wert auch noch von dem unbekannten Werte dieser Funktion in einem gegebenen Punkte ζ abhängig ist.

6. Wollte man unsere Aufgabe auf eine allgemeinere Basis stellen, so würde man zu dem folgendem Problem gelangen:

Gegeben sei eine der Einfachheit halber nebst ihren partiellen Ableitungen bis zu einer hinreichend hohen Ordnung überall stetige Funktion

$$F(x, y, y', \eta, \eta').$$

Es seien weiter zwei Punkte (x_1, y_1), (x_2, y_2) der (x, y)-Ebene und eine der Ungleichheit

$$x_1 < x_0 < x_2$$

genügende feste Zahl x_0 gegeben.

Es sei schließlich $y(x)$ irgendeine, den Randbedingungen

$$y(x_1) = y_1, \quad y(x_2) = y_2$$

genügende, stetige und stetig differentiierbare Funktion.

Wir betrachten das Integral

$$(6) \qquad J[y(x)] = \int_{x_1}^{x_2} F[x, y(x), y'(x), y(x_0), y'(x_0)]\, dx,$$

und wollen es durch geeignete Wahl der Funktion $y(x)$ zu einem Extremum machen.

Das Integral J hängt nicht nur von dem Gesamtverlaufe von $y(x)$ ab, sondern darüber hinaus auch noch von den Werten, welche diese Funktion und ihre Ableitung an einer bestimmten Stelle annimmt.

7. Wir kehren jetzt zu dem in 5. gestellten Problem zurück und bemerken sogleich, daß wir unsere Aufgabe nicht in vollem Umfange gelöst haben. Unser Ergebnis lautet folgenderweise:

Ist

$$\frac{b}{a} \geqq 10$$

und

$$\zeta = b - a,$$

so besitzt das gestellte Problem gewiß keine Lösung im eigentlichen Sinne. Die Differenz

$$V_T(P) - V_T(Q)$$

wird Maximum für

$$\varrho(z) \equiv 0,$$

und zwar ist dann dieses Maximum gleich Null. Wir könnten also von der Existenz der Lösung in diesem Falle nur dann sprechen, wenn wir eine Strecke der z-Achse als eine Rotationsfigur auffassen wollten.

8. Setzt man

$$\zeta = b - a, \quad \varrho(b-a) = \xi,$$

so ist, wie man leicht sieht,

$$V_T(P) - V_T(Q) = 2\pi \int_{-b}^{b} dz \int_0^{\varrho(z)} \frac{t\,dt}{\sqrt{t^2 + (b-z)^2}} - 2\pi \int_{-b}^{b} dz \int_0^{\varrho(z)} \frac{t\,dt}{\sqrt{t^2 + \xi^2 + (b-a-z)^2}}$$

$$- 2\pi \sum_{k=1}^{\infty} \frac{1 \cdot 3 \cdot \ldots \cdot 4k-1}{(2k)!} \cdot \frac{2 \cdot 4 \cdot \ldots \cdot 2k-1}{2 \cdot 4 \cdot \ldots \cdot 2k} \int_{-b}^{b} dz \int_0^{\varrho(z)} \frac{\xi^{2k}\, t^{2k+1}\, dt}{[t^2 + \xi^2 + (b-a-z)^2]^{2k+\frac{1}{2}}},$$

woraus

$$V_T(P) - V_T(Q) \leqq 2\pi \int_{-b}^{+b} dz \int_0^{\varrho(z)} \left\{ \frac{1}{\sqrt{t^2 + (b-z)^2}} - \frac{1}{\sqrt{t^2 + \xi^2 + (b-a-z)^2}} \right\} t\,dt$$

olgt.

Führt man die innere Integration aus, so bekommt man

$$7) \quad V_T(P) - V_T(Q) \leqq 2\pi \int_{-b}^{+b} \left\{ \sqrt{\varrho^2(z) + (b-z)^2} - \sqrt{\varrho^2(z) + \xi^2 + (b-a-z)^2} \right.$$

$$\left. + \sqrt{\xi^2 + (b-a-z)^2} \right\} dz - 4b^2\pi.$$

Wir werden jetzt die rechte Seite durch eine geeignete Wahl von $\varrho(z)$ u einem Maximum zu machen suchen. Schreibt man auf der rechten eite von (7) ausdrücklich

$$\xi = \varrho(b-a),$$

o erkennt man sofort, daß man es mit einem Spezialfalle des in 6. ge-tellten Problems zu tun hat.

Ehe wir weitergehen, wollen wir von der vorausgesetzten Symmetrie-genschaft

$$\varrho(z) = \varrho(-z)$$

Gebrauch machen. Wie man sofort sieht, kann dann (7) wie folgt geschrieben werden:

$$(8)\quad V_T(P)-V_T(Q)\leqq 2\pi\int_0^b\Big\{\sqrt{\varrho^2(z)+(b-z)^2}+\sqrt{\varrho^2(z)+(b+z)^2}$$
$$-\sqrt{\varrho^2(z)+\xi^2+(b-a-z)^2}-\sqrt{\varrho^2(z)+\xi^2+(b-a+z)^2}$$
$$+\sqrt{\xi^2+(b-a-z)^2}+\sqrt{\xi^2+(b-a+z)^2}\Big\}dz-4b^2\pi.$$

Wir werden die Lösung unserer Aufgabe für $b\geqq 10\,a$ auf direktem Wege zu finden versuchen. Zu diesem Zwecke werden wir den unter dem Integralzeichen befindlichen Ausdruck als Funktion von ϱ näher untersuchen. Es genügt natürlich diejenigen Glieder zu nehmen, in welchen ϱ „explizite" vorkommt:

$$(9)\quad \Phi(\varrho,z)=\sqrt{\varrho^2+(b-z)^2}+\sqrt{\varrho^2+(b+z)^2}$$
$$-\sqrt{\varrho^2+\xi^2+(b-a-z)^2}-\sqrt{\varrho^2+\xi^2+(b-a+z)^2}.$$

§ 3.

Eigenschaften der Funktion $\Phi(\varrho,z)$.

9. Die für das Folgende nötigen Eigenschaften der Funktion $\Phi(\varrho,z)$ wollen wir hier in drei Sätzen zusammenfassen:

Wir setzen zur Abkürzung:

$$\frac{1}{\sqrt{\varrho^2+(b-z)^2}}=\beta_1,\quad \frac{1}{\sqrt{\varrho^2+(b+z)^2}}=\beta_2,\quad \frac{1}{\sqrt{\varrho^2+\xi^2+(b-a-z)^2}}=\beta_3,$$
$$\frac{1}{\sqrt{\varrho^2+\xi^2+(b-a+z)^2}}=\beta_4.$$

Es ist dann

$$\Phi(\varrho,z)=\frac{1}{\beta_1}+\frac{1}{\beta_2}-\frac{1}{\beta_3}-\frac{1}{\beta_4},$$

$$\frac{1}{\varrho}\frac{\partial\Phi}{\partial\varrho}=\lambda(\varrho,z)=\beta_1+\beta_2-\beta_3-\beta_4,\qquad \frac{1}{\varrho}\frac{\partial\lambda}{\partial\varrho}=-\beta_1^3-\beta_2^3+\beta_3^3+\beta_4^3,$$

$$\frac{\partial\lambda}{\partial z}=(b-z)\beta_1^3-(b+z)\beta_2^3-(b-a-z)\beta_3^3+(b-a+z)\beta_4^3.$$

Satz 1. *Ist*

$$(10)\qquad 0\leqq\xi\leqq a,\quad 0\leqq\varrho<+\infty,\quad 0\leqq z\leqq b-a,$$

so ist

$$\lambda(\varrho,z)<0.$$

Der Beweis folgt aus der Bemerkung, daß in dem Bereiche (10)

$$\beta_1<\beta_3,\quad \beta_2<\beta_4$$

ist.

Satz 2. *Ist*

$$0 \leqq \xi \leqq a, \quad 0 \leqq \varrho < +\infty, \quad b - a \leqq z < b, \tag{11}$$

so ist

$$\frac{\partial \lambda}{\partial z} > 0.$$

Beweis. In unserem Bereiche gelten die Ungleichheiten

$$\beta_2 < \beta_3, \quad \beta_2 < \beta_4, \quad \beta_2 < \beta_1. \tag{12}$$

Es ist daher

$$\begin{aligned} \frac{\partial \lambda}{\partial z} &> (b-z)\beta_1^3 - (b+z)\beta_2^3 - (b-a-z)\beta_3^3 + (b-a+z)\beta_4^3 \\ &= (b-z)(\beta_1 - \beta_2)(\beta_1^2 + \beta_1\beta_2 + \beta_2^2) > 0. \end{aligned}$$

Satz 3. *Ist*

$$0 \leqq \xi \leqq a \leqq \frac{b}{10}, \quad b - a \leqq z, \quad 0 \leqq \varrho < +\infty,$$

folgt aus

$$\lambda(\varrho, z) = 0 \tag{13}$$

die Ungleichheit

$$\frac{\partial \lambda}{\partial \varrho} < 0.$$

Beweis. Aus $\lambda(\varrho, z) = 0$ folgt

$$\beta_1 = -\beta_2 + \beta_3 + \beta_4$$

und daher wegen (12)

$$-\beta_1^3 - \beta_2^3 + \beta_3^3 + \beta_4^3 = -3(\beta_3 - \beta_2)(\beta_4 - \beta_2)(\beta_3 + \beta_4) < 0.$$

10. Wir betrachten jetzt das durch die Ungleichheiten

$$0 \leqq z \leqq b, \quad 0 \leqq \varrho \leqq b \tag{14}$$

erklärte Rechteck.

Es sind zunächst zwei Fälle möglich:

1. Es ist stets $\lambda(\varrho, z) \neq 0$ und daher $\lambda(\varrho, z) < 0$ (Satz 1).

2. Es gibt Lösungen der Gleichung (13), welche dem Bereiche (14) angehören.

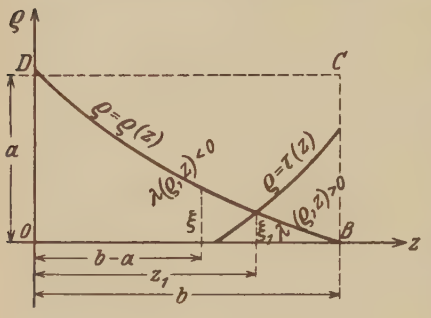
Daß der erste Fall in Wirklichkeit nicht vorkommen kann, ist leicht ersichtlich. Es ist nämlich

$$\lambda(0, z) = \frac{1}{b-z} + \frac{1}{b+z} - \frac{1}{\sqrt{\xi^2 + (b-a-z)^2}} - \frac{1}{\sqrt{\xi^2 + (b-a+z)^2}}.$$

Für kleine $b - z$ ist $\lambda(0, z)$ positiv.

Es bleibt demnach nur der zweite Fall übrig. Wir betrachten die Kurve $\lambda(\varrho, z) = 0$. Sie ist analytisch, und man hat

$$\frac{\partial \lambda}{\partial \varrho} + \frac{\partial \lambda}{\partial z}\frac{dz}{d\varrho} = 0,$$



woraus mit Rücksicht auf die vorher bewiesenen Sätze $\frac{dz}{d\varrho} > 0$ folgt. Die Gleichung unserer Kurve kann also auf die Gestalt

$$\varrho = \tau(z), \tag{15}$$

wo $\tau(z)$ eine eindeutige und wachsende Funktion bezeichnet, gebracht werden.

Unser Rechteck wird durch den Bogen (15) in zwei Bereiche geteilt, in welchen $\lambda(\varrho, z)$ verschiedene Vorzeichen hat (Figur; natürlich wäre es auch denkbar, daß (15) nicht die Seite BC, sondern CD trifft).

Wegen $\lambda = \frac{1}{\varrho}\frac{\partial \Phi}{\partial \varrho}$ ist also $\Phi(\varrho, z)$ als Funktion von ϱ allein in einem dieser Gebiete abnehmend, in dem anderen wachsend.

§ 4.

Reduktion des in § 2 gestellten Problems auf eine Aufgabe der Differentialrechnung.

11. Setzt man vorläufig zur Abkürzung

$$2\pi \int_0^b \left\{\sqrt{\xi^2 + (b-a-z)^2} + \sqrt{\xi^2 + (b-a+z)^2}\right\} dz - 4b^2\pi = R, \tag{16}$$

so ist wegen (8) und (16)

$$V_T(P) - V_T(Q) \leqq 2\pi \int_0^b \Phi\{\varrho(z), z\}\, dz + R. \tag{17}$$

Es sei jetzt $\varrho(z)$ eine beliebige, allen unseren Forderungen genügende Funktion. Die Kurve

$$\varrho = \varrho(z)$$

kann mit der Kurve

$$\varrho = \tau(z)$$

höchstens einen Punkt gemeinsam haben. Wir wollen seine Koordinaten mit (z_1, ξ_1) bezeichnen. Würde es keinen solchen Schnittpunkt geben, so wäre

$$z_1 = b, \quad \xi_1 = \varrho(b)$$

zu setzen (Figur).

Es ist gewiß $z_1 > b - a$.

Wir zerlegen nunmehr das in (17) vorkommende Integral in drei Teile

$$V_T(P) - V_T(Q) \leqq 2\pi \int_0^{b-a} \Phi\{\varrho(z), z\}\, dz + 2\pi \int_{b-a}^{z_1} \Phi\{\varrho(z), z\}\, dz + 2\pi \int_{z_1}^{h} \Phi\{\varrho(z), z\}\, dz + R.$$

Es ist nun für $z \leqq b - a$

$$\varrho(z) \geqq \varrho(b-a) = \xi,$$

und da $\Phi(\varrho, z)$ für diese Werte der Variablen z als Funktion von ϱ abnimmt, so ist

$$\Phi\{\varrho(z), z\} \leqq \Phi(\xi, z).$$

Für $b - a \leqq z \leqq z_1$ ist

$$\varrho(z) \geqq \varrho(z_1) = \xi_1,$$

und da dieser Teil der Kurve $\varrho(z)$ in dem Gebiete $\frac{\partial \Phi}{\partial \varrho} < 0$ enthalten ist, so folgt

$$\Phi\{\varrho(z), z\} \leqq \Phi(\xi_1, z).$$

Für $z_1 \leqq z \leqq b$ ist dagegen

$$\varrho(z) \leqq \varrho(z_1) = \xi_1,$$

der entsprechende Bogen der Kurve $\varrho(z)$ ist in dem Gebiete $\frac{\partial \Phi}{\partial \varrho} > 0$ enthalten; daher gilt

$$\Phi\{\varrho(z), z\} \leqq \Phi(\xi_1, z)$$

auch in diesem Intervall. Es ist demnach

$$(18) \quad V_T(P) - V_T(Q) \leqq 2\pi \int_0^{b-a} \Phi(\xi, z)\, dz + 2\pi \int_{b-a}^{b} \Phi(\xi_1, z)\, dz + R.$$

Der Ausdruck rechter Hand stellt aber eine Funktion der beiden Veränderlichen ξ und ξ_1 dar, und man hat die Aufgabe vor sich, das absolute Maximum dieser Funktion in dem Bereiche

$$(19) \quad 0 \leqq \xi_1 \leqq \xi \leqq a \leqq \frac{b}{10}$$

zu finden.

Das absolute Maximum der rechten Seite der Formel (18) in dem Bereiche (19) wird uns auch eine obere Schranke für $V_T(P) - V_T(Q)$ in $\mathfrak{M}$ liefern, und es wird später äußerst einfach festzustellen sein, daß diese obere Schranke auch die obere Grenze für $V_T(P) - V_T(Q)$ ist.

Wir setzen noch

$$(20) \quad 2\pi \int_0^{b-a} \Phi(\xi, z)\, dz + 2\pi \int_{b-a}^{b} \Phi(\xi_1, z)\, dz + R = \varphi(\xi, \xi_1)$$

und wenden uns der Untersuchung von $\varphi(\xi, \xi_1)$ zu.

§ 5.

Das Maximum von $\varphi(\xi, \xi_1)$.

12. Wir beginnen die Untersuchung von $\varphi(\xi, \xi_1)$ zuerst in bezug auf die Variable ξ_1. Es ist

$$\frac{\partial \varphi}{\partial \xi_1} = 2\pi \int_{b-a}^{b} \frac{\partial \Phi(\xi_1, z)}{\partial \xi_1} dz.$$

Durch direkte Ausrechnung erhält man

$$(21) \qquad \frac{1}{2\pi\xi_1}\frac{\partial \varphi}{\partial \xi_1} = \log \frac{\sqrt{\xi^2+\xi_1^2}}{\xi_1} \cdot \frac{a+\sqrt{\xi_1^2+a^2}}{a+\sqrt{\xi^2+\xi_1^2+a^2}} + \psi_1 + \psi_2,$$

wobei zur Abkürzung

$$22) \quad \psi_1 = \log \frac{2b+\sqrt{\xi_1^2+4b^2}}{2b-a+\sqrt{\xi_1^2+(2b-a)^2}}, \qquad \psi_2 = \log \frac{2(b-a)+\sqrt{\xi^2+\xi_1^2+4(b-a)^2}}{2b-a+\sqrt{\xi^2+\xi_1^2+(2b-a)^2}}$$

gesetzt wurde.

Nun ist ψ_1, als Funktion von b betrachtet, abnehmend, so daß

$$(23) \qquad \psi_1 \geqq \lim_{b\to\infty} \psi_1 = \log 1 = 0$$

gilt.

Demgegenüber ist ψ_2, als Funktion von b betrachtet, wachsend. Es ist also für $b \geqq 4a$

$$(24) \qquad \psi_2 \geqq \log \frac{6a+\sqrt{\xi^2+\xi_1^2+36a^2}}{7a+\sqrt{\xi^2+\xi_1^2+49a^2}}.$$

Aus (21), (22), (23) und (24) folgt

$$\frac{1}{2\pi\xi_1}\frac{\partial \varphi}{\partial \xi_1} \geqq \log \frac{\sqrt{\xi^2+\xi_1^2}}{\xi_1} \cdot \frac{a+\sqrt{\xi_1^2+a^2}}{a+\sqrt{\xi^2+\xi_1^2+a^2}} \cdot \frac{6a+\sqrt{\xi^2+\xi_1^2+36a^2}}{7a+\sqrt{\xi^2+\xi_1^2+49a^2}}.$$

Der Ausdruck $\log \dfrac{a+\sqrt{\xi_1^2+a^2}}{a+\sqrt{\xi^2+\xi_1^2+a^2}}$ ist eine abnehmende Funktion von a, so daß

$$\log \frac{a+\sqrt{\xi_1^2+a^2}}{a+\sqrt{\xi^2+\xi_1^2+a^2}} \geqq \log \frac{\xi+\sqrt{\xi_1^2+\xi^2}}{\xi+\sqrt{\xi_1^2+2\xi^2}}$$

gilt.

Der Ausdruck

$$\log \frac{6a+\sqrt{\xi^2+\xi_1^2+36a^2}}{7a+\sqrt{\xi^2+\xi_1^2+49a^2}}$$

ist dagegen eine wachsende Funktion von a, so daß er sein Maximum für $a \to \infty$ erlangt.

Hieraus folgt also, daß

$$\frac{1}{2\pi\xi_1}\frac{\partial \varphi}{\partial \xi_1} \geqq \log \frac{\sqrt{\xi^2+\xi_1^2}}{\xi_1} \cdot \frac{\xi+\sqrt{\xi_1^2+\xi^2}}{\xi+\sqrt{\xi_1^2+2\xi^2}} \cdot \frac{6}{7}$$

gilt.

Schließlich ist

$$\tau_1(\xi_1) = \log \frac{\sqrt{\xi^2+\xi_1^2}}{\xi_1}$$

eine abnehmende Funktion von ξ_1, so daß

$$\tau_1(\xi_1) \geqq \tau_1(\xi) = \log\sqrt{2}$$

gilt.

Der Ausdruck

$$\tau_2(\xi_1) = \log \frac{\xi+\sqrt{\xi_1^2+\xi^2}}{\xi+\sqrt{\xi_1^2+2\xi^2}}$$

ist eine wachsende Funktion von ξ_1, so daß wir

$$\tau_2(\xi_1) \geqq \tau_2(0) = \log \frac{2}{1+\sqrt{2}}$$

finden.

Es ist demnach

$$\frac{1}{2\pi\xi_1}\frac{\partial\varphi}{\partial\xi_1} \geqq \log\sqrt{2}\cdot\frac{2}{1+\sqrt{2}}\cdot\frac{6}{7} > 0.$$

Hieraus folgt, daß $\varphi(\xi, \xi_1)$ für

$$0 \leqq \xi_1 \leqq \xi \leqq a \leqq \frac{b}{10}$$

eine wachsende Funktion von ξ_1 ist, und daher der Ungleichheit

$$\varphi(\xi, \xi_1) \leqq \varphi(\xi, \xi)$$

genügt.

13. Es ist nun nach (20)

$$\frac{1}{2\pi}\varphi(\xi,\xi) = \int_0^b \Phi(\xi, z)\,dz + \frac{R}{2\pi} \tag{25}$$

$$\begin{aligned}
&= \int_0^b \Big\{\sqrt{\xi^2+(b-z)^2} + \sqrt{\xi^2+(b+z)^2} - \sqrt{2\xi^2+(b-a-z)^2} \\
&\quad - \sqrt{2\xi^2+(b-a+z)^2} + \sqrt{\xi^2+(b-a-z)^2} \\
&\quad + \sqrt{\xi^2+(b-a+z)^2}\Big\}\,dz - 2b^2 \\
&= b\Big\{\sqrt{\xi^2+4b^2} - 2b\Big\} + \frac{a}{2}\Big\{\sqrt{\xi^2+a^2} - \sqrt{2\xi^2+a^2}\Big\} \\
&\quad + \frac{2b-a}{2}\Big\{\sqrt{\xi^2+(2b-a)^2} - \sqrt{2\xi^2+(2b-a)^2}\Big\} \\
&\quad + \frac{\xi^2}{2}\log\frac{2b+\sqrt{\xi^2+4b^2}}{\xi}\cdot\frac{\left(\sqrt{2\xi^2+a^2}-a\right)^2}{\left(2b-a+\sqrt{2\xi^2+(2b-a)^2}\right)^2} \\
&\quad + \frac{\xi^2}{2}\log\frac{2b-a+\sqrt{\xi^2+(2b-a)^2}}{\sqrt{\xi^2+a^2}-a}
\end{aligned}$$

Setzt man noch

$$b = ka, \tag{26}$$

so bekommt man wegen (18), (20) und (25)

$$\frac{V_T(P) - V_T(Q)}{2\pi} \leqq \psi(k, a, \xi)$$

mit

$$\psi(k, a, \xi) = ka\left\{\sqrt{\xi^2 + 4k^2a^2} - 2ka\right\} + \frac{a}{2}\left\{\sqrt{\xi^2 + a^2} - \sqrt{2\xi^2 + a^2}\right\}$$
$$+ \frac{2k-1}{2}a\left\{\sqrt{\xi^2 + (2k-1)^2a^2} - \sqrt{2\xi^2 + (2k-1)^2a^2}\right\}$$
$$+ \frac{\xi^2}{2}\log\frac{2ka + \sqrt{\xi^2 + 4k^2a^2}}{\xi} \cdot \frac{\left(\sqrt{2\xi^2 + a^2} - a\right)^2}{\left\{(2k-1)a + \sqrt{2\xi^2 + (2k-1)^2a^2}\right\}^2}$$
$$+ \frac{\xi^2}{2}\log\frac{(2k-1)a + \sqrt{\xi^2 + (2k-1)^2a^2}}{\sqrt{\xi^2 + a^2} - a}.$$

Es ist nun

$$\frac{\partial\psi}{\partial k} = 2k\left\{\sqrt{\xi^2 + 4k^2a^2} - 2ka\right\} + (2k-1)\left\{\sqrt{\xi^2 + (2k-1)^2a^2}\right.$$
$$\left. - \sqrt{2\xi^2 + (2k-1)^2a^2}\right\} + \left\{\sqrt{a^2 + \xi^2} - \sqrt{a^2 + 2\xi^2}\right\}.$$

Wir betrachten jetzt die Hilfsfunktion

$$\lambda(u) = u\left\{\sqrt{\xi^2 + a^2u^2} - \sqrt{2\xi^2 + a^2u^2}\right\}.$$

Wie man sich leicht überzeugt, nimmt sie monoton ab, so daß für $k > 1$

$$\lambda(2k-1) < \lambda(1)$$

ist, woraus

$$\frac{\partial\psi}{\partial k} \leqq 2k\left\{\sqrt{\xi^2 + 4k^2a^2} - 2ka\right\} + 2\left\{\sqrt{\xi^2 + a^2} - \sqrt{2\xi^2 + a^2}\right\}$$
$$< \frac{\xi^2}{2a} + 2\left\{\sqrt{\xi^2 + a^2} - \sqrt{2\xi^2 + a^2}\right\} < 0$$

folgt.

Wir finden also

$$\psi(k, a, \xi) \leqq \psi(k, \xi, \xi),$$

darum auch

$$\frac{V_T(P) - V_T(Q)}{2\pi\xi^2} \leqq \frac{1}{\xi^2}\psi(k, \xi, \xi) = k\left\{\sqrt{1 + 4k^2} - 2k\right\} \tag{27}$$
$$+ \frac{1}{2}(\sqrt{2} - \sqrt{3}) + \frac{2k-1}{2}\left\{\sqrt{1 + (2k-1)^2} - \sqrt{2 + (2k-1)^2}\right\}$$
$$+ \frac{1}{2}\log\frac{(2k + \sqrt{1+4k^2})(2k - 1 + \sqrt{1 + (2k-1)^2})(\sqrt{3} - 1)^2}{\left\{2k - 1 + \sqrt{2 + (2k-1)^2}\right\}^2(\sqrt{2} - 1)} = r(k).$$

14. Berechnet man $\nu'(k)$, so findet man

$$\frac{1}{2}\nu'(k) = \sqrt{1+4k^2} + \sqrt{1+(2k-1)^2} - \sqrt{2+(2k-1)^2} - 2k,$$

und man überzeugt sich leicht, daß dieser Ausdruck für $k > \frac{1}{2}$ negativ ist.

Für $k > \frac{1}{2}$ ist also $\nu(k)$ abnehmend. Nun findet man durch Ausrechnung, daß

$$\nu(10) = -0{,}0005\ldots$$

ist, so daß für

$$k \geqq 10$$

die Ungleichheit

$$\nu(k) < 0$$

gilt.

Wegen (27) ist also

$$V_T(P) - V_T(Q) < 0,$$

sobald nur

$$b \geqq 10\,a$$

ist.

§ 6.
Schlußfolgerungen.

15. Wir sind mit unserem Beweise zu Ende. Wir haben bewiesen, daß für $b \geqq 10\,a$ immer

$$V_T(P) - V_T(Q) < 0$$

ist.

Es ist nun leicht ersichtlich, daß in unserem Falle die Zahl 0 nicht nur eine obere Schranke, sondern die obere Grenze von $V_T(P) - V_T(Q)$ bildet. Es genügt zu diesem Zwecke eine Folge von T_i zu betrachten, die unseren sämtlichen Bedingungen genügen und deren Volumina gegen Null konvergieren. Die Differenz

$$V_{T_i}(P) - V_{T_i}(Q)$$

konvergiert dann ebenfalls gegen Null.

16. Wir wollen jetzt unser Resultat auf die Gleichgewichtsfiguren rotierender Flüssigkeiten anwenden.

Es sei T eine Gleichgewichtsfigur homogener Flüssigkeit von der Dichte f, die mit der (konstanten) Winkelgeschwindigkeit ω um die z-Achse rotiert. Der Schwerpunkt von T liegt auf der Rotationsachse, und wir können annehmen, daß er auch in der $z = 0$-Ebene liegt. Die $z = 0$-Ebene ist dann die Symmetrieebene von T [6]). Auf dem Rande S von T ist bekanntlich

$$\varkappa f V(P) + \frac{\omega^2}{2}(x^2 + y^2) = \text{konst.} \qquad (\varkappa \text{ bedeutet die Gravitationskonstante}).$$

---

[6]) Vgl. Fußnote [2]).

Wir setzen jetzt weiter voraus, daß unsere Gleichgewichtsfigur T eine Rotationsfigur ist, welche zur Menge $(\mathfrak{M})$ gehört. Unter Benutzung der bisherigen Bezeichnungsweise haben wir

$$\varkappa f V(P) = \varkappa f V(Q) + \frac{\omega^2}{2}\xi^2,$$

woraus

$$V(P) - V(Q) > 0$$

folgt. Demnach muß

$$\frac{b}{a} < 10$$

sein.

(Eingegangen am 5. November 1930.)

Über die Kloostermanschen Summen $S(u, v; q)$.

Von

Hans Salié in Leipzig.

Sei $q \geqq 1$ eine positive ganze Zahl und

$$\varepsilon = e^{\frac{2\pi i}{q}} \tag{1}$$

q-te Einheitswurzel. Sei ferner $(h, q) = 1$, dann gehört zu jedem ganzen $h \geqq 1$ eindeutig ein ganzes $\bar{h}$, so daß

$$h\bar{h} \equiv 1 \pmod{q}, \quad 0 < \bar{h} \leqq q \tag{2}$$

ist.

Bedeuten u, v beliebige ganze Zahlen, so ist nach Herrn Kloosterman [1])

$$S(u, v; q) = \sum_{\substack{0 < h \leqq q \\ (h, q) = 1}}^{'} \varepsilon^{uh + v\bar{h}}. \tag{3}$$

Da die Summen (3) in den Untersuchungen über die Abschätzungen von Fourierkoeffizienten ganzer Modulformen und deren Anwendungen auf die additive Zahlentheorie eine entscheidende Rolle spielen [2]), sei mir gestattet, auf die Eigenschaften der $S(u, v; q)$ zurückzukommen.

Es genügt im folgenden, $S(u, v; q)$ für den Fall

$$q = p^m, \quad m \geqq 1 \text{ ganz}, \quad p \text{ Primzahl} \tag{4}$$

zu untersuchen; denn für jedes

$$q = p_1^{m_1} p_2^{m_2} \dots p_\nu^{m_\nu}, \quad p_1 + p_2 + \dots + p_\nu \tag{5}$$

kann man $S(u, v; q)$ multiplikativ aus Summen (3) mit $q = p_r^{m_r}$ zusammensetzen. Es lassen sich nämlich [3]) bei jedem ganzen u zu jeder ganzen Zahl v

[1]) Acta math. 49 (1926), S. 420.

[2]) Kloosterman, I) Acta math. 49 (1926), S. 407—464; II) Hamb. Abh. V (1927), S. 337—352. — Estermann, Hamb. Abh. VII (1929), S. 82—98 (kurz E. zitiert).

[3]) E., S. 91.

ganze Zahlen $v_1, v_2, \ldots, v_\nu$ finden, so daß

$$S(u, v; q) = \prod_{r=1}^{\nu} S(u, v_r; p_r^{m_r}) \tag{6}$$

wird.

Bisher scheint nicht bemerkt worden zu sein, daß sich für

$$q = p^m, \quad m > 1 \text{ ganz}, \quad p \text{ Primzahl}, \quad (u, q) = (v, q) = 1 \tag{7}$$

$S(u, v; q)$ durch Gaußsche Summen ausdrücken läßt, woraus insbesondere

$$|S(u, v; q)| < C\sqrt{q}, \quad C \text{ unabhängig von } u, v, q, \tag{7a}$$

folgt.

Durch Herrn Estermann[4]) ist bis jetzt für alle q

$$|S(u, v; q)| \leqq \left[d\left(\frac{q}{(u, q)}\right)\right]^{\frac{3}{4}} q^{\frac{3}{4}} (u, q)^{\frac{1}{4}} \tag{8}$$

bekannt, wobei $d(x)$ die Anzahl der positiven Teiler von x bezeichnet.

In der vorliegenden Arbeit werden (§ 1) die $S(u, v; q)$ nach dem quadratischen Restcharakter von u und v eingeteilt. Die dadurch sich ergebenden Summen $f(\varepsilon^\nu)$ und $g(\varepsilon^\nu, N_0)$ — N_0 ein Nichtrest mod q — erweisen sich in manchen Formeln zweckmäßiger als die bezüglich u, v nicht getrennten $S(u, v; q)$. Vor allem tritt die Berechtigung der Aufspaltung beim Beweise von (7a) zutage (§ 2, § 3).

Wenn q Primzahl ist (§§ 4, 5, 6), gelingt zwar eine explizite Auswertung von $f(\varepsilon^\nu)$ und $g(\varepsilon^\nu)$ durch Gaußsche Summen nicht, wie es unter der Voraussetzung von (7) möglich war. Doch lassen sich für die $f(\varepsilon^\nu)$ Formeln aufstellen, die mit Gaußschen Summen verknüpft und der Bauart nach verwandt erscheinen.

Weiter zeigt sich, daß $f(\varepsilon^\nu)^2$ und $g(\varepsilon^\nu)^2$ linear darstellbar (mit Koeffizienten $\leqq 1$) durch die $f(\varepsilon^\nu)$ und $g(\varepsilon^\nu)$ sind.

Schließlich geben die mit $f(\varepsilon^\nu)$ und $g(\varepsilon^\nu)$ gebildeten Summen Zusammenhänge mit Formeln, wie sie in der Theorie der Verteilung der quadratischen Reste aufgestellt werden[5]).

---

[4]) E., S. 83. Vermutlich ist die Abschätzung (7a) in der Form

$$|S(u, v; q)| < C q^{\frac{1}{2}+\varepsilon} (u, q)^{\frac{1}{2}} \quad \varepsilon > 0 \text{ beliebig}$$

für alle q gültig. Auf den Beweis dieser Abschätzung hoffe ich demnächst zurückkommen zu können. Der Hauptsatz in K. II (siehe auch E., S. 83) würde dann

$$c_n = O\left(n^{\frac{\varkappa}{2}-\frac{1}{4}+\varepsilon}\right)$$

lauten.

[5]) E. Jacobsthal, J. f. d. r. u. angew. Math. **132** (1907), S. 238—245. — H. Hopf, Math. Zeitschr. **32** (1930), S. 222—231.

§ 1.

Alle $S(u, v; q)$ sind reell, wie aus der Darstellung

$$
\begin{aligned}
(9) \qquad S(u, v; q) &= 1 && \text{für } q = 1\\
&= (-1)^{u+v} && \text{für } q = 2\\
&= \sum_{\substack{0<h\leqq \frac{q}{2}\\ (h,q)=1}} \left(\varepsilon^{uh+v\bar{h}} + \varepsilon^{-(uh+v\bar{h})}\right) && \text{für } q \geqq 3\\
&= 2 \sum_{\substack{0<h\leqq \frac{q}{2}\\ (h,q)=1}} \cos \frac{2\pi}{q}(uh + v\bar{h})
\end{aligned}
$$

hervorgeht, die überdies

$$(10) \qquad S(-u, -v; q) = S(u, v; q)$$

zeigt.

Unter den bekannten, für alle q gültigen Eigenschaften der $S(u, v; q)$ hebe ich hervor:

1. die Symmetrie-Eigenschaft:

$$(11) \qquad S(u, v; q) = S(v, u; q)\ ^{6});$$

2. $(u, q) = 1$, v beliebig,

$$(12) \qquad S(u, v; q) = S(1, uv; q)\ ^{7})$$

und entsprechend

$(v, q) = 1$, u beliebig,

$$(12a) \qquad S(u, v; q) = S(1, uv; q);$$

3. $v \equiv 0 \pmod q$

$$(13) \qquad S(u, v; q) = \sum_{t/(u,q)} t\,\mu\left(\frac{q}{t}\right).$$

4. $(u, q) = 1$, $v \equiv 0 \pmod q$

$$(13a) \qquad S(u, v; q) = \mu(q)$$

und $u \equiv 0$, $v \equiv 0 \pmod q$

$$(13b) \qquad S(u, v; q) = \varphi(q),$$

wobei $\mu(q)$ das Möbiussche Symbol und $\varphi(q)$ die Eulersche Funktion bezeichnet.

---

[6]) E., S. 89.
[7]) E., S. 87.

Sei nunmehr

$$f(\varepsilon) = \sum_{\substack{0<h<q \\ (h,q)=1}} \varepsilon^{h+\bar{h}}, \qquad q \geqq 2, \tag{14}$$

und

$$g(\varepsilon, N_0) = \sum_{\substack{0<h<q \\ (h,q)=1}} \varepsilon^{h+N_0\bar{h}}, \qquad q \geqq 3, \tag{15}$$

mit $(N_0, q) = 1$, N_0 quadratischer Nichtrest mod q.

Bildet man für jedes zu q prime ν $f(\varepsilon^\nu)$ und $g(\varepsilon^\nu, N_0)$ unter Berücksichtigung von (12), so folgt:

$$f(\varepsilon^\nu) = \sum_{\substack{0<h<q \\ (h,q)=1}} \varepsilon^{\nu(h+\bar{h})} = \sum_{\substack{0<h<q \\ (h,q)=1}} \varepsilon^{h+\nu^2\bar{h}} = S(1, \nu^2; q), \tag{16}$$

$$g(\varepsilon^\nu, N_0) = \sum_{\substack{0<h<q \\ (h,q)=1}} \varepsilon^{\nu(h+N_0\bar{h})} = \sum_{\substack{0<h<q \\ (h,q)=1}} \varepsilon^{h+N_0\nu^2\bar{h}} = S(1, N_0\nu^2; q). \tag{17}$$

Es ist also:

$$(u, q) = 1, \ (v, q) = 1,$$

$$S(u, v; q) = \begin{cases} f(\varepsilon^\nu), & \nu^2 \equiv uv \pmod q \\ g(\varepsilon^\nu, N_0), & N_0\nu^2 \equiv uv \pmod q, \end{cases} \tag{18}$$

je nachdem sich ν aus der ersten oder zweiten Kongruenz finden läßt[8]).

Offenbar ist

$$f(\varepsilon^{-\nu}) = f(\varepsilon^\nu), \quad g(\varepsilon^{-\nu}, N_0) = g(\varepsilon^\nu, N_0). \tag{19}$$

Die unter (12) bis (18) aufgeführten Gleichungen setzen voraus, daß mindestens eine der Zahlen u, v zu q teilerfremd ist. Wenn u und v mit q einen Teiler gemeinsam haben, sind Reduktionen der $S(u, v; q)$ möglich, die jedoch nur für $q = p^m$, $m \geqq 1$ durchgeführt werden sollen.

§ 2.

In diesem Abschnitt ist beständig

$$q = p^m, \quad m \geqq 2, \quad p \text{ Primzahl}. \tag{20}$$

Die Zahlen

$$0 < h \leqq q, \quad (h, q) = 1$$

stimmen überein mit den Zahlen

$$h = l + rp^{m-1}, \quad \begin{matrix} r = 0, 1, 2, \ldots, p-1, \\ l = 1, 2, 3, \ldots, p^{m-1}-1, \end{matrix} \qquad l \not\equiv 0 \pmod p;$$

[8]) Im allgemeinen nur bei passend gewähltem N_0.

zu denen

$$\bar{h} \equiv \bar{l} - r\bar{l}^2 p^{m-1} \pmod{q}, \qquad m \geqq 2$$

gehören, wie man durch $h\bar{h} \equiv 1 \pmod{q}$ bestätigt.

Hierdurch läßt sich (3) in der Form

$$S(u, v; q) = \sum_{\substack{0<l<p^{m-1} \\ l \not\equiv 0 \pmod{p}}} \varepsilon^{ul+v\bar{l}} \cdot \sum_{r=0}^{p-1} e^{\frac{2\pi i}{p}(u-v\bar{l}^2)r} \tag{21}$$

schreiben.

Die Summe über r ist stets Null, wenn $u \equiv 0 \pmod{p}$ und $v \not\equiv 0 \pmod{p}$ ist, also auch

$$S(u, v; q) = 0 \text{ [9]}, \qquad u \equiv 0 \pmod{p}, \qquad v \not\equiv 0 \pmod{p}, \tag{22}$$

und wegen (11) (oder direkt aus (21))

$$S(u, v; q) = 0, \qquad u \not\equiv 0 \pmod{p}, \qquad v \equiv 0 \pmod{p}. \tag{22a}$$

Sind u und v beide durch p teilbar, so ist nach Gl. (21)

$$S(u, v; q) = p \sum_{\substack{0<l<p^{m-1} \\ l \not\equiv 0 \pmod{p}}} \varepsilon^{ul+v\bar{l}} = pS\left(\frac{u}{p}, \frac{v}{p}; \frac{q}{p}\right),$$

oder durch wiederholte Anwendung dieser Beziehung:

$$S(u, v; q) = p^k S\left(\frac{u}{p^k}, \frac{v}{p^k}; p^{m-k}\right) \text{ [10]}, \qquad \begin{matrix} u \equiv 0 \pmod{p^k}, \\ v \equiv 0 \pmod{p^k}, \\ 1 \leqq k < m \end{matrix} \tag{23}$$

und

$$S(u, v; q) = p^m - p^{m-1}, \qquad u \equiv v \equiv 0 \pmod{q}. \tag{23a}$$

Es zerfallen die zu

$$u \not\equiv 0 \pmod{p}, \qquad v \not\equiv 0 \pmod{p}$$

gehörigen $S(u, v; q)$ nach (18) für $p > 2$ in $\frac{1}{2}\varphi(q)$ Funktionen $f(\varepsilon^\nu)$ und $\frac{1}{2}\varphi(q)$ Funktionen $g(\varepsilon^\nu, N_0)$, wobei

$$\nu \not\equiv 0 \pmod{p}, \qquad 1 \leqq \nu \leqq \frac{q-1}{2}$$

genommen werden kann, da (19) gilt. N_0 sei ein fester Nichtrest $\not\equiv 0 \pmod{p}$. Bei Übergang zu einem anderen Nichtrest $N_0' \not\equiv 0 \pmod{p}$ ergeben sich dieselben Funktionen $g(\varepsilon^\nu, N_0)$ abgesehen von der Reihenfolge.

Nach (21) ist

$$g(\varepsilon^\nu, N_0) = \sum_{\substack{0<l<p^{m-1} \\ l \not\equiv 0 \pmod{p}}} \varepsilon^{\nu(l+N_0\bar{l})} \cdot \sum_{r=0}^{p-1} e^{\frac{2\pi i}{p} r(1-N_0\bar{l}^2)\nu} \tag{24}$$

---

[9]) E., S. 89.

[10]) E., S. 90.

und

$$(25)\qquad f(\varepsilon^\nu)=\sum_{\substack{0<l<p^{m-1}\\ l\not\equiv 0\,(\mathrm{mod}\,p)}}\varepsilon^{\nu(l+\bar l)}\sum_{r=0}^{p-1}e^{\frac{2\pi i}{p}r(1-\bar l^2)\nu}$$

oder, nach Auswertung der Summen über r:

$$(26)\qquad g(\varepsilon^\nu, N_0)=0,\quad p>2,\quad m\geqq 2,$$

$$(27)\qquad f(\varepsilon^\nu)=p\sum_{\substack{0<l<p^{m-1}\\ l\equiv\pm 1\,(\mathrm{mod}\,p)}}\varepsilon^{\nu(l+\bar l)},\quad p>2,\quad m\geqq 2.$$

(27) ist nur der Spezialfall $\tau=1$ der folgenden Gleichung

$$(28)\qquad f(\varepsilon^\nu)=p^\tau\sum_{\substack{0<h<p^{m-\tau}\\ h\equiv\pm 1\,(\mathrm{mod}\,p^\tau)}}\varepsilon^{\nu(h+\bar h)},\quad m\geqq 2\tau,\quad p>2,$$

die durch Induktionsschluß von τ auf $\tau+1$ für alle τ als gültig erwiesen wird.

Wenn nämlich $m\geqq 2\tau+2$ ist, darf man in (28) die durch

$$0<h<p^{m-\tau},\qquad h\equiv\pm 1\,(\mathrm{mod}\,p^\tau)$$

gekennzeichneten Zahlen aufspalten in:

$$(29)\qquad h=l+rp^{m-\tau-1},\qquad \begin{matrix} l=1,\ldots,p^{m-\tau-1}-1,\\ r=0,1,2,\ldots,p-1,\end{matrix}\qquad l\equiv\pm 1\,(\mathrm{mod}\,p^\tau)$$

und $\bar h$ in der Form schreiben:

$$\bar h\equiv\bar l-r\bar l^2p^{m-\tau-1}.$$

Durch Einsetzen von (29) in (28) entsteht:

$$(30)\qquad f(\varepsilon^\nu)=p^\tau\sum_{\substack{0<l<p^{m-\tau-1}\\ l\equiv\pm 1\,(\mathrm{mod}\,p^\tau)}}\varepsilon^{\nu(l+\bar l)}\cdot\sum_{r=0}^{p-1}\varepsilon^{\nu(1-\bar l^2)p^{m-\tau-1}r},\qquad m\geqq 2(\tau+1)$$

oder unter Beachtung von

$$\sum_{\substack{r=0\\ l\equiv\pm 1\,(\mathrm{mod}\,p^\tau)}}^{p-1}\varepsilon^{\nu(1-\bar l^2)p^{m-\tau-1}r}=\begin{cases}p, & l\equiv\pm 1\,(\mathrm{mod}\,p^{\tau+1}),\\ 0 & \text{sonst},\end{cases}$$

$$(31)\qquad f(\varepsilon^\nu)=p^{\tau+1}\sum_{\substack{0<l<p^{m-\tau-1}\\ l\equiv\pm 1\,(\mathrm{mod}\,p^{\tau+1})}}\varepsilon^{\nu(l+\bar l)},\qquad m\geqq 2(\tau+1),\quad p>2,$$

wodurch (28) bewiesen ist.

Für gerades m ergibt $\tau = \frac{m}{2}$ in (28) eingesetzt, wegen $h = 1$, $p^{\frac{m}{2}} - 1$ mit $\bar{h} \equiv 1$, $-p^{\frac{m}{2}} - 1 \pmod q$

$$(32) \qquad f(\varepsilon^\nu) = p^{\frac{m}{2}}(\varepsilon^{2\nu} + \varepsilon^{-2\nu}) \qquad p > 2, \quad m \equiv 0 \pmod 2$$
$$= 2\sqrt{q}\cos\frac{4\pi\nu}{q}. \qquad \nu \not\equiv 0 \pmod p.$$

Entsprechend entsteht bei ungeradem m für $\tau = \frac{m-1}{2}$ mit $h = \pm 1 + rp^{\frac{m-1}{2}}$ und $\bar{h} \equiv \pm 1 - rp^{\frac{m-1}{2}} \pm r^2 p^{m-1} \pmod q$ — $r \to$ volles Restsystem mod p —

$$f(\varepsilon^\nu) = p^{\frac{m-1}{2}}\left[\varepsilon^{2\nu}\sum_{r=0}^{p-1} e^{\frac{2\pi i}{p}\nu r^2} + \varepsilon^{-2\nu}\sum_{r=0}^{p-1} e^{-\frac{2\pi i}{p}\nu r^2}\right]$$
$$= \left(\frac{\nu}{p}\right) p^{\frac{m-1}{2}}\sqrt{(-1)^{\frac{p-1}{2}} p}\left(\varepsilon^{2\nu} + (-1)^{\frac{p-1}{2}}\varepsilon^{-2\nu}\right),$$

d. h. es ist für $m \equiv 1 \pmod 2$, $p > 2$

$$(33) \qquad f(\varepsilon^\nu) = \begin{cases} 2\left(\frac{\nu}{q}\right)\sqrt{q}\cos\frac{4\pi\nu}{q}, & p \equiv 1 \pmod 4 \\ -2\left(\frac{\nu}{q}\right)\sqrt{q}\sin\frac{4\pi\nu}{q}, & p \equiv 3 \pmod 4. \end{cases}$$

Zusammenfassend läßt sich sagen, daß im Falle

$$q = p^m, \quad p \geqq 3, \quad \text{Primzahl}, \quad m \geqq 2$$

nur die $S(u, v; q)$ von Null verschieden sind, für die entweder

$$u \equiv 0 \pmod q, \quad v \equiv 0 \pmod q \quad \text{(vgl. (24a))},$$

oder

$$(u, q) = (v, q) + p^{m-1}\ ^{11}),$$

$$\frac{u}{(u, q)} \not\equiv 0 \left(\mathrm{mod}\ \frac{q}{(u, q)}\right), \quad \frac{v}{(u, q)} \not\equiv 0 \left(\mathrm{mod}\ \frac{q}{(v, q)}\right),$$

$$\frac{uv}{(uv, q)} \text{ quadratischer Rest mod } \frac{q}{(u, q)}$$

ist. Hinsichtlich der Größenordnung gilt für alle $S(u, v; q)$:

$$|S(u, v; q)| < 2q^{\frac{1}{2}}(u, q)^{\frac{1}{2}}, \quad (u, v, q) + p^{m-1}.$$

---

[11]) Für $(u, q) = (v, q) = p^{m-1}$ ist $S(u, v; q) = p^{m-1} S\left(\frac{u}{p^{m-1}}, \frac{v}{p^{m-1}}; p\right)$.

§ 3.

Der Fall $q = 2^m$ erfordert besondere Behandlung:

$$(34)\qquad \begin{aligned} m &= 1 \quad S(u, v; q) = (-1)^{u+v}, \\ m &= 2 \quad S(u, v; q) = i^{u+v} + i^{-(u+v)}. \end{aligned}$$

Für $m \geqq 3$ gibt es $\frac{1}{4}\varphi(q) = 2^{m-3}$ Funktionen $f(\varepsilon^\nu)$ und $\frac{3}{4}\varphi(q) = 3 \cdot 2^{m-3}$ Funktionen $g(\varepsilon^\nu, N_0)$. Sobald $m \geqq 6$ ist, sind auch hier alle

$$(35)\qquad g(\varepsilon^\nu, N_0) = 0.$$

Beweis. Ist

$$(36)\qquad h = k + 2^{m-\lambda}, \quad m \geqq 2\lambda, \quad \lambda \geqq 0 \text{ ganz}, \quad k \text{ ungerade},$$

so folgt

$$\bar{h} \equiv \bar{k} - \bar{k}^3 2^{m-\lambda} \pmod{q},$$

oder, falls $\lambda = 1, 2, 3$ genommen wird:

$$\bar{h} \equiv \bar{k} - 2^{m-\lambda} \pmod{q} \qquad (\lambda = 1, 2, 3),$$

da $\bar{k}^2 \equiv 1 \pmod 8$ ist.

Setzt man nun (36) mit $\lambda = 1$ in

$$(37)\qquad S(1, \mu; 2^m) = \sum_{\substack{1 \leqq h < 2^m \\ h \equiv 1 \pmod 2}} \varepsilon^{h + \mu\bar{h}}, \quad \mu \text{ beliebig, ganz},$$

ein, so entsteht

$$(38)\qquad S(1, \mu; 2^m) = (1 + (-1)^{\mu+1}) \sum_{\substack{1 \leqq h < 2^{m-1} \\ h \equiv 1 \pmod 2}} \varepsilon^{h + \mu\bar{h}}, \quad m \geqq 2.$$

Für gerade μ ist $S(1, \mu; 2^m) = 0$, $m \geqq 2$, was oben schon § 2 zu entnehmen ist. Sei nunmehr μ ungerade. Wendet man (36) jetzt mit $\lambda = 2$ auf (38) an, so wird:

$$(39)\qquad S(1, \mu; 2^m) = 2(1 + i^{1-\mu}) \sum_{\substack{1 \leqq h < 2^{m-2} \\ h \equiv 1 \pmod 2}} \varepsilon^{h + \mu\bar{h}}, \quad m \geqq 4,$$

woraus

$$(39a)\qquad S(1, \mu; 2^m) = 0, \quad \mu \equiv 3 \pmod 4, \quad m \geqq 4$$

hervorgeht.

Schließlich werde für $\mu = 4\alpha + 1$ in (39) die Substitution (36) mit $\lambda = 3$ eingeführt, um

$$(40)\quad S(1, \mu; 2^m) = 4\left(1 + (-1)^{\frac{\mu-1}{4}}\right) \sum_{\substack{1 \leqq h < 2^{m-3} \\ h \equiv 1 \pmod 2}} \varepsilon^{h + \mu\bar{h}}, \quad m \geqq 6, \quad \mu \equiv 1 \pmod 4$$

u erhalten, d. h.

40a) $$S(1, \mu; 2^m) = 0, \qquad \mu \equiv 5 \pmod 8, \quad m \geqq 6,$$

40b) $$S(1, \mu; 2^m) = 8 \sum_{\substack{1 \leqq h < 2^{m-2} \\ h \equiv 1 \pmod 2}} \varepsilon^{h+\mu\bar{h}}, \qquad \mu \equiv 1 \pmod 8, \quad m \geqq 6.$$

Durch (39a) und (40a) ist (35) bewiesen. Aus (18) und (40b) folgt:

41) $$f(\varepsilon^\nu) = 8 \sum_{\substack{0 < h < 2^{m-2} \\ h \equiv 1 \pmod 2}} \varepsilon^{\nu(h+\bar{h})}, \qquad m \geqq 6, \quad \nu \equiv 1 \pmod 2.$$

Für die weitere Berechnung von $f(\varepsilon^\nu)$ brauche ich den folgenden Hilfssatz. Ist

42) $$m \geqq 2\tau, \qquad h = 2^{m-\tau} - k, \qquad k^2 \equiv 1 \pmod{2^{\tau-1}},$$

o ist

42a) $$\bar{h} \equiv -2^{m-\tau} - \bar{k} + \varrho_k 2^{m-1} \pmod{2^m},$$

obei

$$\varrho_k = \begin{cases} 0 & \text{für } k^2 \equiv 1 \pmod{2^\tau} \\ 1 & \text{sonst} \end{cases}$$

t.

Beweis. Es ist

$$h\bar{h} \equiv -2^{2m-2\tau} + 2^{m-\tau}(k - \bar{k}) + 1 + \varrho_k 2^{2m-\tau-1} - k\varrho_k 2^{m-1}$$
$$\equiv 1 + 2^{m-\tau}(k - \bar{k}) - \varrho_k \cdot 2^{m-1},$$

a $m \geqq 2\tau$ und $k \equiv 1 \pmod 2$ ist.

Nun folgt aus $k^2 \equiv 1 \pmod{2^{\tau-1}}$

$$k - \bar{k} \equiv \varrho_k 2^{\tau-1} \pmod{2^\tau},$$

o daß also $h \cdot \bar{h} \equiv 1 \pmod{2^m}$ und damit der Ausdruck (42a) als richtig achgewiesen ist.

Wird der obige Hilfssatz mit $\tau = 3$ für alle ungeraden h

$$2^{m-4} < h < 2^{m-3}$$

ngewendet (ϱ_k ist dann stets Null) und in (41) verwertet, so entsteht:

43) $$f(\varepsilon^\nu) = 8 \sum_{\substack{0 < h < 2^{m-4} \\ h \equiv \pm 1 \pmod 4}} (\varepsilon^{\nu(h+\bar{h})} + \varepsilon^{-\nu(h+\bar{h})}), \qquad m \geqq 6$$

der, wenn jetzt $\tau = 4$ bedeutet:

44) $$f(\varepsilon^\nu) = 2^{\tau-1} \sum_{\substack{0 < h < 2^{m-\tau} \\ h \equiv \pm 1 \pmod{2^{\tau-2}}}} (\varepsilon^{\nu(h+\bar{h})} + \varepsilon^{-\nu(h+\bar{h})}), \qquad m \geqq 2(\tau - 1).$$

Ich zeige durch vollständige Induktion, daß (44) für alle τ, $\leqq \tau \leqq \frac{m}{2} + 1$ gilt.

In der zweiten Summe der aus (44) folgenden Gleichung

$$(44a)\qquad f(\varepsilon^\nu) = 2^{\tau-1} \sum_{\substack{0<h\leqq 2^{m-\tau-1} \\ h\equiv \pm 1 \,(\mathrm{mod}\, 2^{\tau-2})}} (\varepsilon^{\nu(h+\bar{h})} + \varepsilon^{-\nu(h+\bar{h})}) + 2^{\tau-1} \sum_{\substack{2^{m-\tau-1}<h<2^{m-\tau} \\ h\equiv \pm 1 \,(\mathrm{mod}\, 2^{\tau-2})}} (\varepsilon^{\nu(h+\bar{h})} + \varepsilon^{-\nu(h+\bar{h})})$$

werde unter der Voraussetzung $m \geqq 2\tau$ (42) eingeführt, so daß

$$(45)\qquad f(\varepsilon^\nu) = 2^{\tau-1} \sum_{\substack{0<h<2^{m-\tau-1} \\ h\equiv \pm 1 \,(\mathrm{mod}\, 2^{\tau-2})}} [\varepsilon^{\nu(h+\bar{h})} + \varepsilon^{-\nu(h+\bar{h})} + (-1)^{\varrho_h}(\varepsilon^{\nu(h+\bar{h})} + \varepsilon^{-\nu(h+\bar{h})})], \qquad m \geqq 2\tau$$

wird, oder unter Benutzung der Bedeutung der ϱ_h

$$(46)\qquad f(\varepsilon^\nu) = 2^{\tau} \sum_{\substack{0<h<2^{m-\tau-1} \\ h\equiv \pm 1 \,(\mathrm{mod}\, 2^{\tau-1})}} (\varepsilon^{\nu(h+\bar{h})} + \varepsilon^{-\nu(h+\bar{h})}), \qquad m \geqq 2\tau$$

eine Gleichung, die aus (44) durch Übergang von τ auf $\tau+1$ zu erhalten ist. Die Angabe der endgültigen Formeln für $f(\varepsilon^\nu)$ aus (46) macht nun keine Schwierigkeiten mehr:

a) m gerade, $\tau = \frac{m}{2}$, es ist $h = 1$ und $h = 2^{\frac{m}{2}-1} - 1$, wozu $\bar{h} \equiv 1$ bzw. $\bar{h} \equiv -2^{m-2} - 2^{\frac{m}{2}-1} - 1$ gehören;

$$(47)\qquad f(\varepsilon^\nu) = 2^{\frac{m}{2}}(\varepsilon^{2\nu} + \varepsilon^{-2\nu} + \varepsilon^{2\nu+2^{m-2}\nu} + \varepsilon^{-2\nu-2^{m-2}\nu}) = 2\sqrt{2q}\cos\left(\frac{\pi}{4} + (-1)^{\frac{\nu-1}{2}}\frac{4\pi\nu}{q}\right), \qquad \begin{matrix} m \geqq 6,\ \nu \equiv 1 \bmod 2, \\ m \equiv 0 \bmod 2. \end{matrix}$$

b) m ungerade, $\tau = \frac{m-1}{2}$, $m \geqq 9$

$$h = 1,\quad 2^{\frac{m-3}{2}} - 1,\quad 2^{\frac{m-3}{2}} + 1,\quad 2^{\frac{m-1}{2}} - 1$$

$$\bar{h} \equiv 1,\quad -2^{m-3} - 2^{\frac{m-3}{2}} - 1,\quad 2^{m-3} - 2^{\frac{m-3}{2}} + 1,\quad -2^{m-1} - 2^{\frac{m-1}{2}} - 1,$$

$$(48)\qquad f(\varepsilon^\nu) = 2^{\frac{m-1}{2}}(\varepsilon^{2\nu} + \varepsilon^{-2\nu} + 2\varepsilon^{2\nu+2^{m-2}\nu} + 2\varepsilon^{-2\nu-2^{m-2}\nu} + \varepsilon^{2\nu+2^{m-1}\nu} + \varepsilon^{-2\nu-2^{m-1}\nu})$$

$$= 2\sqrt{2q}\cos\left(\frac{4\pi\nu}{q} + \frac{\pi\nu}{4}\right), \quad \nu \equiv 1\,(\mathrm{mod}\,2),\ m \equiv 1\,(\mathrm{mod}\,2),\ m \geqq 9$$

$$= 2(-1)^{\frac{\nu^2-1}{8}}\sqrt{2q}\cos\left(\frac{\pi}{4} + (-1)^{\frac{\nu-1}{2}}\frac{4\pi\nu}{q}\right).$$

Von den noch fehlenden $f(\varepsilon^\nu)$ stellt man leicht fest:

$$m = 3, \quad f(\varepsilon^\nu) = 0$$
$$m = 4, \quad f(\varepsilon^\nu) = 4(-1)^{\frac{\nu^2-1}{8}}\sqrt{2}$$
$$m = 5, \quad f(\varepsilon^\nu) = 0$$
$$m = 7, \quad f(\varepsilon^\nu) = -32\cos\left(\frac{4\pi\nu}{2^7} + \frac{\pi\nu}{4}\right).$$

Es ist stets

$$|f(\varepsilon^\nu)| \leqq 2\sqrt{2q} \quad \text{für alle } q = 2^m. \tag{49}$$

§ 4.

Es sei q ungerade Primzahl. Zur Übersicht werden die möglichen Werte der $S(u, v; q)$ zusammengestellt:

$$\begin{cases} u \equiv 0, \ v \equiv 0, & S(u, v; q) = q - 1; \\ u \equiv 0, \ v \not\equiv 0, & S(u, v; q) = -1; \\ u \not\equiv 0, \ v \equiv 0, & S(u, v; q) = -1; \\ u \not\equiv 0, \ v \not\equiv 0, & S(u, v; q) = \begin{cases} f(\varepsilon^\nu), & \nu^2 \equiv uv \pmod q; \\ g(\varepsilon^\nu, N_0), & \nu^2 \equiv uv\,\bar{N}_0 \pmod q. \end{cases} \end{cases} \tag{50}$$

Da im folgenden der einmal gewählte Nichtrest N_0 beibehalten wird, soll $g(\varepsilon^\nu)$ statt $g(\varepsilon^\nu, N_0)$ geschrieben werden.

Mit Hilfe des Legendreschen Symbols

$$\left(\frac{a}{q}\right) = \left\{ \begin{array}{ll} 0, & a \equiv 0 \pmod q, \\ 1, & a \text{ quadr. Rest} \\ -1, & a \text{ quadr. Nichtrest} \end{array} \right\} \bmod q$$

kann man (14) und (15) in die Form

$$f(\varepsilon^\nu) = \sum_{h=0}^{q-1} \left(\frac{h^2-4}{q}\right) \varepsilon^{\nu h}, \tag{51}$$

$$g(\varepsilon^\nu) = \sum_{h=0}^{q-1} \left(\frac{h^2-4N_0}{q}\right) \varepsilon^{\nu h} \tag{52}$$

bringen, da für jede eindeutige Funktion $F(x)$ mit der Periode q die Beziehung[12])

$$\lambda \not\equiv 0 \pmod q, \quad \sum_{h=1}^{q-1} F(h + \lambda\bar{h}) = \sum_{h=0}^{q-1} F(h) + \sum_{h=0}^{q-1} \left(\frac{h^2-4\lambda}{q}\right) F(h) \tag{53}$$

[12]) E. Jacobsthal, a. a. O. [5]) S. 239.

gilt. Für $\nu \equiv 0 \pmod q$ soll

$$f(\varepsilon^\nu) = g(\varepsilon^\nu) = -1, \quad \nu \equiv 0 \pmod q \tag{52a}$$

festgesetzt werden in Übereinstimmung mit

$$\sum_{h=0}^{q-1}\left(\frac{h^2-4}{q}\right) = \sum_{h=0}^{q-1}\left(\frac{h^2-4N_0}{q}\right) = -1.$$

Die Beziehung (53) werde jetzt noch zweimal angewendet:

1. für $\lambda = 1$, $F(x) = \varepsilon^{\nu x^2}$ liefert sie

$$\varepsilon^{2\nu}\sum_{h=1}^{q-1}\varepsilon^{\nu(h^2+\bar h^2)} = \left(\frac{\nu}{q}\right)\sqrt{(-1)^{\frac{q-1}{2}}q} + \sum_{h=0}^{q-1}\left(\frac{h^2-4}{q}\right)\varepsilon^{\nu h^2}; \tag{53a}$$

2. für $\lambda = N_0$, $F(x) = \varepsilon^{\nu \bar N_0 x^2}$ liefert sie

$$\varepsilon^{2\nu}\sum_{h=1}^{q-1}\varepsilon^{\nu(\bar N_0 h^2 + N_0 \bar h^2)} = -\left(\frac{\nu}{q}\right)\sqrt{(-1)^{\frac{q-1}{2}}q} + \sum_{h=0}^{q-1}\left(\frac{h^2-4N_0}{q}\right)\varepsilon^{\nu \bar N_0 h^2}. \tag{53b}$$

Wegen

$$\sum_{h=0}^{q-1}\left(\frac{h^2-4}{q}\right)\varepsilon^{\nu h^2} - \sum_{h=0}^{q-1}\left(\frac{h^2-4N_0}{q}\right)\varepsilon^{\nu \bar N_0 h^2} = 2\sum_{k=0}^{q-1}\left(\frac{k-4}{q}\right)\varepsilon^{\nu k}$$

$$= 2\varepsilon^{4\nu}\left(\frac{\nu}{q}\right)\sqrt{(-1)^{\frac{q-1}{2}}q}$$

erhält man durch Subtraktion und Addition von (53a) und (53b)

$$\sum_{h=1}^{q-1}\left(\frac{h}{q}\right)\varepsilon^{(h+\bar h)\nu} = \left(\frac{\nu}{q}\right)\sqrt{(-1)^{\frac{q-1}{2}}q}\,(\varepsilon^{2\nu} + \varepsilon^{-2\nu}), \tag{54}$$

$$f(\varepsilon^\nu) = -\left(\frac{\nu}{q}\right)\varepsilon^{2\nu}\sqrt{(-1)^{\frac{q-1}{2}}q} + \varepsilon^{-2\nu}\sum_{h=0}^{q-1}\left(\frac{h^2-4}{q}\right)\varepsilon^{\nu h^2} \tag{55}$$

bzw.

$$f(\varepsilon^\nu) = \left(\frac{\nu}{q}\right)\varepsilon^{2\nu}\sqrt{(-1)^{\frac{q-1}{2}}q} + \varepsilon^{-2\nu}\sum_{h=0}^{q-1}\left(\frac{h^2-4N_0}{q}\right)\varepsilon^{\nu \bar N_0 h^2}. \tag{56}$$

Für $g(\varepsilon^\nu)$ bestehen keine genau entsprechenden Formeln:

$$\sum_{h=1}^{q-1}\left(\frac{h}{q}\right)\varepsilon^{(h+N_0\bar h)\nu} = 0; \tag{57}$$

denn die Summe links nimmt ihren negativen Wert an, wenn $h \equiv N_0 \bar k$ eingeführt wird. Aus (57) folgt:

$$g(\varepsilon^\nu) = \sum_{h=1}^{q-1}\varepsilon^{\nu(lh^2 + N_0\bar h^2 \bar l)}, \quad l \not\equiv 0 \pmod q. \tag{58}$$

Ist χ ein Charakter mod q, so ist

$$\sum_{\chi} \chi(lh) = \begin{cases} q-1, & lh \equiv 1 \pmod{q}, \\ 0 & \text{sonst} \end{cases}$$

woraus

$$\sum_{l=1}^{q-1} \varepsilon^{vl} \sum_{\chi} \chi(l)\,\chi(h) = (q-1)\,\varepsilon^{v\bar{h}}$$

und

$$S(u, v; q) = \frac{1}{q-1} \sum_{\chi} \sum_{h=1}^{q-1} \chi(h)\,\varepsilon^{uh} \cdot \sum_{h=1}^{q-1} \chi(h)\,\varepsilon^{vh} \tag{59}$$

folgt. Diese Form der Kloostermanschen Summen liefert für $u \equiv v \equiv \nu \pmod{q}$

$$f(\varepsilon^{\nu}) = \frac{1}{q-1} \sum_{\chi} \left(\sum_{h=1}^{q-1} \chi(h)\,\varepsilon^{\nu h} \right)^2, \qquad \nu \not\equiv 0 \pmod{q}. \tag{59a}$$

Hierin ist bekanntlich[18])

$$\left| \sum_{h=1}^{q-1} \chi(h)\,\varepsilon^{\nu h} \right| \leqq \sqrt{q}.$$

Das Auftreten von $\sqrt{(-1)^{\frac{q-1}{2}} q}$ in den verschiedenen Darstellungen legt die Bildung von $f(\varepsilon^{\nu})^2$ und $g(\varepsilon^{\nu})^2$ nahe.

Nach (58) ist:

$$\varepsilon^{\nu(l+N_0\bar{l})}\, g(\varepsilon^{\nu}) = \sum_{h=1}^{q-1} \varepsilon^{\nu(lh^2+N_0\bar{l}\bar{h}^2+l+N_0\bar{l})},$$

$$g(\varepsilon^{\nu})^2 = \sum_{h=1}^{q-1} \sum_{l=1}^{q-1} \varepsilon^{\nu(lh^2+N_0\bar{l}\bar{h}^2+l+N_0\bar{l})},$$

und, wenn $l \equiv k\bar{h}$ eingeführt wird,

$$g(\varepsilon^{\nu})^2 = \sum_{h=1}^{q-1} \sum_{k=1}^{q-1} \varepsilon^{\nu(h+\bar{h})(k+N_0\bar{k})},$$

also

$$g(\varepsilon^{\nu})^2 = \left(1 + (-1)^{\frac{q-1}{2}}\right) q + \sum_{h=1}^{q-1} g(\varepsilon^{\nu(h+\bar{h})}), \qquad \nu \not\equiv 0 \pmod{q} \tag{60}$$

und

$$g(\varepsilon^{\nu})^2 = \left(1 - (-1)^{\frac{q-1}{2}}\right) q + \sum_{h=1}^{q-1} f(\varepsilon^{\nu(h+N_0\bar{h})}), \qquad \nu \not\equiv 0 \pmod{q} \tag{60a}$$

unter Berücksichtigung der Festsetzung (52a).

---

[18]) E. Landau, Vorlesungen über Zahlentheorie (Leipzig 1927), Bd. I, S. 189.

Schließlich ist nach (53)

$$(61)\qquad g(\varepsilon^\nu)^2 = q + \sum_{h=0}^{q-1}\left(\frac{h^2-4}{q}\right) g(\varepsilon^{\nu h}) = q + \sum_{h=0}^{q-1}\left(\frac{h^2-4N_0}{q}\right) f(\varepsilon^{\nu h}),$$

$$\nu \not\equiv 0 \pmod q,$$

wenn man

$$(62)\qquad \sum_{\nu=0}^{q-1} f(\varepsilon^\nu) = -\sum_{\nu=0}^{q-1} g(\varepsilon^\nu) = (-1)^{\frac{q-1}{2}} q$$

beachtet.

Ähnlich beweist man:

$$(63)\quad f(\varepsilon^\nu)^2 = \left(1+(-1)^{\frac{q-1}{2}}\right) q - (-1)^{\frac{q-1}{2}} q(\varepsilon^{2\nu}+\varepsilon^{-2\nu})^2 + \sum_{h=0}^{q-1} f(\varepsilon^{\nu(h+\bar{h})})$$

$$= \left(1-(-1)^{\frac{q-1}{2}}\right) q + (-1)^{\frac{q-1}{2}} q(\varepsilon^{2\nu}+\varepsilon^{-2\nu})^2 + \sum_{h=0}^{q-1} g(\varepsilon^{\nu(h+\bar{N}_0\bar{h})})$$

und

$$(63\text{a})\quad f(\varepsilon^\nu)^2 = q - (-1)^{\frac{q-1}{2}} q(\varepsilon^{4\nu}+\varepsilon^{-4\nu}) + \sum_{h=0}^{q-1}\left(\frac{h^2-4}{q}\right) f(\varepsilon^{\nu h})$$

$$= q + (-1)^{\frac{q-1}{2}} q(\varepsilon^{4\nu}+\varepsilon^{-4\nu}) + \sum_{h=0}^{q-1}\left(\frac{h^2-4\bar{N}_0}{q}\right) g(\varepsilon^{\nu h}).$$

An den gegebenen Darstellungen für $f(\varepsilon^\nu)^2$ und $g(\varepsilon^\nu)^2$ ist bemerkenswert, daß sie sich wieder in gewissen $f(\varepsilon^\nu)$ und $g(\varepsilon^\nu)$ ausdrücken.

§ 5.

Da[14])

$$(64)\qquad \operatorname{Max} S(1,\lambda;q)^2 = \lim_{n\to\infty} \sqrt[n]{\sum_{\lambda=0}^{q-1} S(1,\lambda;q)^{2n}}$$

ist, kann die Ermittlung der genauen Größenordnung der $S(u,v;q)$ in q, die für q als ungerade Primzahl noch aussteht, auf die Bestimmung der Potenzsummen der $S(u,v;q)$ zurückgeführt werden.

Um zu einem Ausdruck für

$$\sum_{\lambda=0}^{q-1} S(1,\lambda:q)^n = \frac{1}{2}\sum_{\nu=0}^{q-1} f(\varepsilon^\nu)^n + \frac{1}{2}\sum_{\nu=0}^{q-1} g(\varepsilon^\nu)^n$$

---

[14]) Pólya-Szegö, Aufgaben und Lehrsätze aus der Analysis (Berlin 1925), Bd. I, S. 77, Nr. 195.

zu gelangen, setze man in

$$\sum_{\lambda=0}^{q-1} \varepsilon^{-\lambda k} S(1, \lambda; q) = q \varepsilon^{\bar{k}}, \quad k \not\equiv 0 \pmod q$$

ein:

$$k \equiv -(k_1 + k_2 + \ldots + k_{n-1}) \pmod q, \quad n \geqq 2, \quad 1 \leqq k_\nu \leqq q-1$$
$$(\nu = 1, 2, \ldots, n-1),$$

multipliziere mit $\varepsilon^{\bar{k}_1 + \bar{k}_2 + \ldots + \bar{k}_{n-1}}$ und summiere über alle k_ν. Es entsteht:

$$\sum_{\lambda=0}^{q-1} S(1, \lambda; q)^n = q \sum_{k_1, k_2, \ldots, k_{n-1}}^{1,2,\ldots,q-1} \varepsilon^{\bar{k}_1 + \bar{k}_2 + \ldots + \bar{k}_{n-1} - \overline{(k_1 + k_2 + \ldots + k_{n-1})}}$$

$$= q \sum_{\nu=1}^{q-1} \sum_{\substack{k_1, k_2, \ldots, k_{n-1} \\ k_1 + k_2 + \ldots + k_{n-1} \equiv \nu \pmod q}} \varepsilon^{\bar{k}_1 + \bar{k}_2 + \ldots + \bar{k}_{n-1} - \bar{\nu}}$$

$$= q \sum_{\nu=1}^{q-1} \sum_{\substack{l_1, l_2, \ldots, l_{n-1} \\ l_1 + l_2 + \ldots + l_{n-1} \equiv 1 \pmod q}} \varepsilon^{\bar{\nu}(\bar{l}_1 + \bar{l}_2 + \ldots + \bar{l}_{n-1} - 1)},$$

falls $k_r \equiv \nu l_r \pmod q$ $(r = 1, 2, \ldots, n-1)$ gesetzt wird.

Nach Ausführung der Summation über ν ergibt sich:

$$\sum_{\lambda=0}^{q-1} S(1, \lambda; q)^n = q(q-1) M_{n-1} - q M_{n-1}^*, \tag{65}$$

wobei M_{n-1} die Anzahl der Lösungen $1 \leqq l_r \leqq q-1$ $(r = 1, 2, \ldots, n-1)$ von

$$\left| \begin{array}{l} l_1 + l_2 + \ldots + l_{n-1} \equiv 1 \\ \bar{l}_1 + \bar{l}_2 + \ldots + \bar{l}_{n-1} \equiv 1 \end{array} \pmod q \right| \tag{66}$$

und M_{n-1}^* die Anzahl der Lösungen $1 \leqq l_r \leqq q-1$ $(r = 1, 2, \ldots, n-1)$ von

$$\left| \begin{array}{l} l_1 + l_2 + \ldots + l_{n-1} \equiv 1 \\ \bar{l}_1 + \bar{l}_2 + \ldots + \bar{l}_{n-1} \not\equiv 1 \end{array} \pmod q \right| \tag{67}$$

bezeichnet.

Durch eine leichte Abzählung findet man, daß

$$M_{n-1} + M_{n-1}^* = \frac{(q-1)^{n-1} - (-1)^{n-1}}{q} \tag{68}$$

ist als Anzahl der Zahlen $1 \leqq l_r \leqq q-1$ $(r = 1, 2, \ldots, n-1)$, für die

$$l_1 + l_2 + \ldots + l_{n-1} \equiv 1 \pmod q.$$

(65) geht durch (68) über in:

$$\sum_{\lambda=0}^{q-1} S(1, \lambda; q)^n = q^2 M_{n-1} - (q-1)^{n-1} + (-1)^{n-1} \qquad (n \geqq 2). \tag{69}$$

Für $n = 2, 3, 4$ gelingt die Bestimmung von M_{n-1} sehr leicht. Es ist

$$M_1 = 1,$$

$$M_2 = \begin{cases} 1, & q = 3, \\ 1 + \left(\frac{-3}{q}\right), & q > 3, \end{cases}$$

denn $\begin{vmatrix} l_1 + l_2 \equiv 1 \\ \bar{l}_1 + \bar{l}_2 \equiv 1 \end{vmatrix}$ ist für $q > 3$ nur im Falle $\left(\frac{-3}{q}\right) = +1$ lösbar.

Bei festem $l_3 \not\equiv 1$ hat (66) für $n = 4$ die Lösungen

$$\begin{matrix} l_1 \equiv 1, & l_2 \equiv -l_3 \\ l_1 \equiv -l_3, & l_2 \equiv 1 \end{matrix} \pmod q$$

und

Die beiden Lösungspaare fallen nur für $l_3 \equiv q - 1$ zusammen. Ist $l_3 \equiv 1$, so sind $(q-1)$ Lösungspaare

$$l_1 \equiv l, \qquad l_2 \equiv -l \qquad (l = 1, 2, \ldots, q-1)$$

vorhanden. Somit ist

$$M_3 = 2(q-3) + 1 + (q-1) = 3(q-2)$$

und[15])

$$(70) \qquad \begin{cases} \sum_{\lambda=0}^{q-1} S(1, \lambda; q)^2 = q^2 - q. \\ \sum_{\lambda=0}^{q-1} S(1, \lambda; q)^3 = \left(\frac{-3}{q}\right) q^2 + 2q, \\ \sum_{\lambda=0}^{q-1} S(1, \lambda; q)^4 = 2q^3 - 3q^2 - 3q. \end{cases}$$

Der Limes (64) ist hinsichtlich seines Verhaltens für große q bis jetzt nicht bekannt. (70) liefert jedoch, da

$$\left(\frac{\sum_{\lambda=0}^{q-1} S(1, \lambda; q)^{2n}}{q} \right)^{\frac{1}{n}}$$

eine monoton wachsende Funktion von n ist[16]),

$$(71) \qquad \sum_{\lambda=0}^{q-1} S(1, \lambda; q)^{2n} > 2^{\frac{n}{2}} (q-1)^{n+1},$$ [17])

und außerdem

$$(72) \qquad |S(1, \lambda; q)| < \sqrt[4]{\sum_{\lambda=0}^{q-1} S(1, \lambda; q)^4} < 2^{\frac{1}{4}} q^{\frac{3}{4}},$$

---

[15]) Bei E., S. 89 findet sich nur $\sum_{\substack{0 < w \leqq q \\ (w, q) = 1}} |S(1, w; q)|^4 \leqq 8q^3$ abgeschätzt.

[16]) a. a. O. [14]) S. 54, Nr. 82.

[17]) Vermutlich ist auch $\sum_{\lambda=0}^{q-1} S(1, \lambda; q)^{2n} < (Cq)^{n+1}$, wo C von q unabhängig ist.

eine Abschätzung, die sich nur in der Konstanten von der Abschätzung (8) des Herrn Estermann unterscheidet.

Zu einer zweiten Darstellung der Summen (69) gelangt man, wenn man die Gleichungen

$$(73)\qquad \sum_{\nu=0}^{q-1} f(\varepsilon^\nu)^n = q \sum_{\substack{0 \leqq h_\nu < q \\ h_1+\ldots+h_n \equiv 0 \,(\mathrm{mod}\, q)}} \left(\frac{h_1^2-\alpha^2}{q}\right)\left(\frac{h_2^2-\alpha^2}{q}\right)\cdots\left(\frac{h_n^2-\alpha^2}{q}\right)$$

$$\alpha \not\equiv 0 \,(\mathrm{mod}\, q), \qquad n \geqq 1, \text{ ganz}$$

$$(74)\qquad \sum_{\nu=0}^{q-1} g(\varepsilon^\nu)^n = q \sum_{\substack{0 \leqq h_\nu < q \\ h_1+\ldots+h_n \equiv 0 \,(\mathrm{mod}\, q)}} \left(\frac{h_1^2-\alpha^2 N_0}{q}\right)\left(\frac{h_2^2-\alpha^2 N_0}{q}\right)\cdots\left(\frac{h_n^2-\alpha^2 N_0}{q}\right),$$

die sich unmittelbar aus (51) und (52) mit $\alpha = 2$ ergeben — offenbar aber für alle $\alpha \not\equiv 0 \,(\mathrm{mod}\, q)$ gelten — zu

$$(75)\qquad \sum_{\lambda=0}^{q-1} S(1, \lambda; q)^n = \frac{q}{q-1} \sum_{a=1}^{q-1} \sum_{\substack{0 \leqq h_\nu < q \\ h_1+\ldots+h_n \equiv 0 \,(\mathrm{mod}\, q)}} \left(\frac{h_1^2+a}{q}\right)\left(\frac{h_2^2+a}{q}\right)\cdots\left(\frac{h_n^2+a}{q}\right)$$

zusammenzieht. $n = 2$ liefert in (73) und (74)

$$\sum_{\nu=0}^{q-1} f(\varepsilon^\nu)^2 = q^2 - 2q,$$

$$\sum_{\nu=0}^{q-1} g(\varepsilon^\nu)^2 = q^2,$$

woraus hervorgeht, daß in einer für alle ν gültigen asymptotischen Abschätzung

$$\left.\begin{matrix} f(\varepsilon^\nu) \\ g(\varepsilon^\nu) \end{matrix}\right\} = O(q^\beta),$$

$\beta \geqq \frac{1}{2}$ sein muß (nach (72) ist $\beta \leqq \frac{3}{4}$).

Es sollen im folgenden noch einige Summen, in denen $f(\varepsilon^\nu)$ und $g(\varepsilon^\nu)$ auftreten, gebildet werden, die wie (73) und (74) Zusammenhänge mit Summen Legendrescher Symbole ergeben.

Man stellt leicht fest, daß z. B.

$$(76)\qquad \sum_{\nu=0}^{q-1} \left(\frac{\nu}{q}\right) \varepsilon^{k\nu} f(\varepsilon^\nu) = \sqrt{(-1)^{\frac{q-1}{2}} q} \sum_{h=0}^{q-1} \left(\frac{h^2-4}{q}\right)\left(\frac{h+k}{q}\right),$$

$$(77)\qquad \sum_{\nu=0}^{q-1} f(\varepsilon^\nu) f(\varepsilon^{\lambda\nu}) = q \sum_{h=0}^{q-1} \left(\frac{h^2-4}{q}\right)\left(\frac{h^2-4\lambda^2}{q}\right),$$

$$(78)\qquad \sum_{\nu=0}^{q-1} \varepsilon^{4\lambda\nu} f(\varepsilon^\nu)^2 = q \sum_{h=0}^{q-1} \left(\frac{h^2-4}{q}\right)\left(\frac{(h-4\lambda)^2-4}{q}\right).$$

Die drei Summen, deren asymptotisches Verhalten in q bei Fragen der Verteilung der quadratischen Reste eine Rolle spielt[18]), hängen untereinander einfach zusammen. Ihre wahre Größenordnung ist im allgemeinen noch nicht bekannt, dagegen ist sie im Spezialfall $k = 0$ in (76) durch Jacobsthal[19])

$$\left[\sum_{h=1}^{q-1}\left(\frac{h+\bar{h}}{q}\right)\right]^2+\left[\sum_{h=1}^{q-1}\left(\frac{h+N_0\bar{h}}{q}\right)\right]^2=\begin{cases}0, & q\equiv 3\,(\mathrm{mod}\,4)\\ 4q, & q\equiv 1\,(\mathrm{mod}\,4)\end{cases}$$

bestimmt, die der Beziehung

$$\left[\sum_{\nu=0}^{q-1}\left(\frac{\nu}{q}\right)f(\varepsilon^\nu)\right]^2+\left[\sum_{\nu=0}^{q-1}\left(\frac{\nu}{q}\right)g(\varepsilon^\nu)\right]^2=\begin{cases}0, & q\equiv 3\,(\mathrm{mod}\,4)\\ 4q^2, & q\equiv 1\,(\mathrm{mod}\,4)\end{cases}$$

entspricht. Auf weitere (Gleichungen (76) bis (78)) ähnliche Formeln, die sich in großer Zahl finden lassen, soll nicht eingegangen werden. Vielleicht kann für die Ermittlung der Größenordnung der Summen Legendrescher Symbole die Kenntnis des Zusammenhangs mit den Kloostermanschen Summen von Nutzen sein.

§ 6.

Zum Schluß sei noch darauf hingewiesen, daß $f(\varepsilon^\nu)$ und $g(\varepsilon^\nu)$ primitive Zahlen des durch Adjunktion von $\varepsilon+\varepsilon^{-1}$ zum natürlichen Rationalitätsbereich $\mathfrak{R}$ erweiterten Körpers vom Grade $\frac{q-1}{2}$ sind, wenn q eine ungerade Primzahl ist. Es genügen also $f(\varepsilon^\nu)$ und $g(\varepsilon^\nu)$ je einer in $\mathfrak{R}$ irreduziblen Gleichung $\frac{q-1}{2}$-ten Grades.

Da nach einem bekannten Satz über zyklische Determinanten[20])

$$\prod_{\nu=0}^{q-1}(x-f(\varepsilon^\nu))=-\begin{vmatrix} a_0-x & a_1 & \dots & a_{q-1}\\ a_{q-1} & a_0-x & \dots & a_{q-2}\\ \dots & \dots & \dots & \dots\\ a_1 & a_2 & \dots & a_0-x\end{vmatrix},\qquad a_h=\left(\frac{h^2-4}{q}\right),$$

[18]) H. Hopf, a. a. O. [5]). Der Zusammenhang mit dem dortigen $\chi(c)$ wird etwa durch die Formel

$$\sum_{h=0}^{q-1}\left(\frac{h^2+a}{q}\right)\left(\frac{h^2+ca}{q}\right)=-1+\left(\frac{a}{q}\right)\sum_{h=0}^{q-1}\left(\frac{h}{q}\right)\left(\frac{h+1}{q}\right)\left(\frac{h+c}{q}\right),\qquad a\not\equiv 0\,(\mathrm{mod}\,q)$$

hergestellt.

[19]) a. a. O. [5]) S. 239—240.

[20]) G. Kowalewski, Einführung in die Determinantentheorie (Leipzig 1909), S. 116.

oder unter beiderseitiger Abspaltung des Faktors

$$x - f(\varepsilon^0) = x + 1 = -(a_0 + a_1 + \ldots + a_{q-1} - x):$$

$$F(x)^2 \equiv \left[\prod_{\nu=1}^{\frac{q-1}{2}} (x - f(\varepsilon^\nu))\right]^2 = \begin{vmatrix} 1 & a_1 & \ldots & a_{q-1} \\ 1 & a_0 - x & \ldots & a_{q-2} \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ 1 & a_2 & \ldots & a_0 - x \end{vmatrix}$$

ist[21]), stellt

$$F(x) = 0$$

die irreduzible Gleichung dar, der die $f(\varepsilon^\nu)$ genügen.

Entsprechend findet man die irreduzible Gleichung für die $g(\varepsilon^\nu)$, wenn $a_h = \left(\frac{h^2 - 4N_0}{q}\right)$ bedeutet.

---

[21]) Nach dem Hadamardschen Determinantensatz ist übrigens

$$\prod_{\nu=1}^{\frac{q-1}{2}} |f(\varepsilon^\nu)| \leqq q^{\frac{q}{4}}.$$

(Eingegangen am 10. Januar 1931.)

Stetige Funktionen in topologischen Räumen.

Ernst Zermelo zum 60. Geburtstag am 27. Juli 1931 gewidmet.

Von

Reinhold Baer in Halle a. S. und Friedrich Levi in Leipzig.

Einleitung.

Versteht man unter einer stetigen Funktion in einem topologischen Raume[1]) $\mathfrak{T}$ eine stetige Abbildung von $\mathfrak{T}$ auf einen topologischen Raum $\mathfrak{T}^*$, so wird eine solche Funktion α wesentlich durch die von ihr erzeugte Einteilung des Raumes $\mathfrak{T}$ in Kongruenzklassen bestimmt. [Dabei heißen zwei Punkte kongruent, wenn ihnen durch die stetige Funktion α derselbe Funktionswert zugeordnet wird.] Man kann nämlich leicht einsehen, daß alle stetigen Funktionen, die dieselbe Einteilung in Kongruenzklassen bewirken, stetige Funktionen von einer unter ihnen sind. Macht man eine geeignete Einschränkung über die Topologie in den Bildräumen $\mathfrak{T}^*$ [Räume minimaler Dichte], so werden diese Funktionen durch die Einteilung in Kongruenzklassen sogar bis auf topologische Abbildungen eindeutig bestimmt (§ 1).

Man wird aber hierüber hinaus bestrebt sein, die Werträume der stetigen Funktionen in $\mathfrak{T}$ nicht nur durch Zerlegung von $\mathfrak{T}$ zu gewinnen, sondern wird auch versuchen, Repräsentantensysteme der Kongruenzklassen aufzufinden, die als Werträume dienen können. Bekannte Beispiele der Funktionentheorie [z. B. die elliptischen Funktionen] zeigen, daß dies schon in sehr einfachen Fällen nicht mehr möglich ist. Nach dem Muster der Funktionentheorie wird man sich also auf die Aufsuchung von Diskontinuitätsbereichen beschränken, das sind in $\mathfrak{T}$ offene Mengen, auf denen die Funktion topologisch ist, mit der Eigenschaft, daß der Bildraum $\mathfrak{T}^*$ von $\mathfrak{T}$ gerade die abgeschlossene Hülle ihres Bildbereiches ist. Dann und nur dann ist jeder Punkt von $\mathfrak{T}$ in einem Diskontinuitätsbereich enthalten, wenn die Funktion im kleinen topologisch ist (§ 2).

[1]) Im Sinne von F. Hausdorff, Mengenlehre, 1. Aufl., Leipzig 1914, S. 213. Vgl. auch R. Baer, Beziehungen zwischen den Grundbegriffen der Topologie, Sitz.-Ber. d. Heidelberger Akad. d. Wiss., Math.-nat. Kl., 1929 (15).

Versteht man unter der Überdeckungszahl die kleinste Zahl offener, punktfremder Mengen, auf denen die stetige Funktion α topologisch ist, die ausreicht, um $\mathfrak{T}$ bis auf eine nirgends dichte Menge zu überdecken [solche Überdeckungssysteme gibt es immer, wenn die Funktion im kleinen topologisch ist], und unter der Blätterzahl die obere Grenze der Mächtigkeiten der einzelnen Kongruenzklassen, so sind diese beiden Zahlen gleich, wenn eine der beiden Zahlen endlich und die Funktion im kleinen topologisch ist. Für den Fall, daß beide Zahlen unendlich sind, tritt an die Stelle der Gleichheit eine Ungleichung (§§ 3, 4).

Die durch eine stetige Funktion definierte Kongruenzrelation steht in einer gewissen Analogie zu der Äquivalenzrelation, die durch eine Gruppe T topologischer Abbildungen von $\mathfrak{T}$ auf sich erzeugt wird[2]). Will man diese beiden Relationen zueinander in Beziehung setzen, so wird man die folgenden drei Stufen zu unterscheiden haben:

1. Die Kongruenzklassen gehen bei T wieder in Kongruenzklassen über: zulässige Funktionen, die eine Verpflanzung von T auf den Bildraum gestatten (§ 5).

2. Die Kongruenzklassen sind bezgl. T invariant: automorphe Funktionen. Dann gibt es stets eine im wesentlichen eindeutig bestimmte automorphe Funktion φ, so daß die Gesamtheit der automorphen Funktionen mit der Gesamtheit der stetigen Funktionen von φ identisch ist (§ 6).

3. Kongruenzklassen und Äquivalenzklassen fallen zusammen. Hier wird insbesondere der Fall untersucht, daß die stetige Funktion im kleinen topologisch ist: diskontinuierliche Gruppe. Ist dann $\mathfrak{T}$ kompakt und im kleinen zusammenhängend, so ist jede diskontinuierliche Gruppe endlich (§ 7).

Wesentlich benutzt wird [besonders in den §§ 2 bis 4] der Zermelosche Wohlordnungssatz.

Bezeichnungen.

Umgebung eines Punktes $\mathfrak{p}$ [bezw. einer Menge $\mathfrak{M}$] ist jede den Punkt $\mathfrak{p}$ [bezw. die Menge $\mathfrak{M}$] enthaltende offene Menge[3]).

Hp = Häufungspunkt.
$\mathfrak{A} \cap \mathfrak{B}$ = Durchschnitt von $\mathfrak{A}$ und $\mathfrak{B}$.
$\mathfrak{A} \vee \mathfrak{B}$ = Vereinigung von $\mathfrak{A}$ und $\mathfrak{B}$.
$\mathfrak{A} < \mathfrak{B}$ = $\mathfrak{A}$ ist echte Teilmenge von $\mathfrak{B}$.
$\mathfrak{B} - \mathfrak{A}$ = Komplement der Teilmenge $\mathfrak{A}$ von $\mathfrak{B}$ in $\mathfrak{B}$.
$\overline{\mathfrak{A}}$ = abgeschlossene Hülle von $\mathfrak{A}$.
$\mathfrak{A} = 0$ heißt: die Menge $\mathfrak{A}$ ist leer.

---

[2]) Vgl. hierzu R. Baer und F. Levi, Ränder topologischer Räume, Ber. d. math.-phys. Kl. d. Sächsischen Akad. d. Wiss. zu Leipzig, 82 (1930), S. 171—189.

[3]) Mit H. Tietze, Beiträge zur allgemeinen Topologie I, Math. Annalen 88 (1923), S. 290—312.

§ 1.

Stetige Funktionen und stetige Zerlegungen.

Eine in $\mathfrak{T}$ definierte Funktion α ist eine eindeutige Abbildung des Raumes $\mathfrak{T}$ auf einen Bildraum $\mathfrak{T}^* = \alpha(\mathfrak{T})$. Punkte, denen durch α derselbe Bildpunkt zugeordnet wird, heißen *kongruent* [bezgl. α]. Jede Funktion erzeugt also eine Zerlegung von $\mathfrak{T}$ in Kongruenzklassen

$$\{\mathfrak{C}_\sigma\}: \mathfrak{T} = \mathfrak{C}_1 \vee \mathfrak{C}_2 \vee \ldots \vee \mathfrak{C}_\sigma \vee \ldots \quad \text{mit} \quad \mathfrak{C}_\nu \cap \mathfrak{C}_\mu = 0 \quad \text{für} \quad \nu \neq \mu.$$

Die Punkte aus einem $\mathfrak{C}_\sigma$ sind also die sämtlichen Urbilder eines und nur eines Punktes aus dem Bildraum $\mathfrak{T}^*$.

Eine Funktion α heißt *stetig* in $\mathfrak{T}$, wenn das Urbild einer jeden in $\alpha(\mathfrak{T})$ offenen Menge in $\mathfrak{T}$ offen ist[4]). Ist dann $\mathfrak{M}^*$ in $\mathfrak{T}^* = \alpha(\mathfrak{T})$ abgeschlossen und $\mathfrak{M}$ das Urbild von $\mathfrak{M}^*$ bezgl. der stetigen Funktion α, so ist $\mathfrak{T}^* - \mathfrak{M}^*$ offen und ebenso das Urbild $\mathfrak{T} - \mathfrak{M}$. Also ist das Urbild $\mathfrak{M}$ einer abgeschlossenen Menge auch abgeschlossen.

Die von einer stetigen Funktion erzeugte Zerlegung in Kongruenzklassen heißt *stetige Zerlegung*[5]).

Durch eine stetige Zerlegung wird der Bildraum der sie definierenden stetigen Abbildung nicht einmal bis auf topologische Abbildungen eindeutig bestimmt[6]). Zu einem Eindeutigkeitssatz kommt man aber, wenn man sich auf Funktionen beschränkt, welche $\mathfrak{T}$ auf Bildräume von minimaler Dichte abbilden. Dabei heiße $\alpha(\mathfrak{T})$ bezgl. α *von minimaler Dichte*, wenn jede Menge in $\alpha(\mathfrak{T})$ offen ist, deren [vollständiges] Urbild in $\mathfrak{T}$ offen ist.

Satz 1. a) *Eine Zerlegung von $\mathfrak{T}$ in Mengen $\mathfrak{C}_\nu$ ist dann und nur dann stetig, wenn es zu irgend zwei verschiedenen Komponenten der Zerlegung punktfremde Umgebungen gibt, die aus vollständigen Kongruenzklassen bestehen.*

b) *Ist α stetig, so wird $\mathfrak{T}^* = \alpha(\mathfrak{T})$ dadurch zu einem topologischen Raum $\mathfrak{T}^{**}$, der bezgl. α von minimaler Dichte ist, daß man alle und nur die Mengen aus $\mathfrak{T}^*$, deren Urbilder offen sind, zu offenen Mengen erklärt*[7]). *Die zugehörige Abbildung von $\mathfrak{T}^{**}$ auf $\mathfrak{T}^*$ ist eineindeutig und stetig.*

---

[4]) In Übereinstimmung mit dem Sprachgebrauch; vgl. etwa F. Hausdorff, a. a. O. S. 361, Satz III. — Dagegen wird gelegentlich auch unter einer stetigen Funktion eine solche verstanden, die offene Mengen auf offene Mengen abbildet; vgl. z. B. S. Lefschetz, Topology, New York 1930, S. 3.

[5]) Vgl. P. Alexandroff, Über stetige Abbildungen kompakter Räume, Math. Annalen 96 (1926), S. 555—571. — Die dort auf S. 557 gegebene Definition 2 deckt sich im wesentlichen mit der unsrigen, wie aus unserem Satz 1 folgt.

[6]) Dies zeigt z. B. das von Alexandroff a. a. O. S. 558—559 angegebene Beispiel.

[7]) Also ist jede in $\mathfrak{T}^*$ offene Menge auch in $\mathfrak{T}^{**}$ offen oder, was dasselbe sagt: Ist $\mathfrak{p}^*$ in $\mathfrak{T}^{**}$ Hp der Teilmenge $\mathfrak{M}^*$, so auch in $\mathfrak{T}^*$.

c) *Ist α_i für $i=1, 2$ stetig in $\mathfrak{T}$, $\alpha_i(\mathfrak{T})$ von minimaler Dichte bezgl. α_i und erzeugen α_1 und α_2 dieselbe Zerlegung von $\mathfrak{T}$ in Kongruenzklassen, so ist die Zuordnung*

$$\alpha_1(\mathfrak{p}) \leftrightarrow \alpha_2(\mathfrak{p})$$

eine topologische Abbildung der Bildräume $\alpha_1(\mathfrak{T})$ und $\alpha_2(\mathfrak{T})$ aufeinander.

d) *Ist α stetig in $\mathfrak{T}$, so sind die Kongruenzklassen nach α in $\mathfrak{T}$ abgeschlossen.*

Beweis. ad a) Daß die Bedingung notwendig ist, folgt einerseits daraus, daß in $\alpha(\mathfrak{T})$ das Trennbarkeitsaxiom erfüllt ist und andererseits die Urbilder offener Mengen aus $\alpha(\mathfrak{T})$ in $\mathfrak{T}$ offen sind.

Es sei also umgekehrt $\{\mathfrak{C}_\sigma\}$ eine Zerlegung von $\mathfrak{T}$, die der Bedingung von Satz 1. a) genügt. Um zu zeigen, daß sie die Zerlegung in Kongruenzklassen nach einer geeigneten, in $\mathfrak{T}$ stetigen Funktion α ist, konstruiere man den *aus dieser Zerlegung abgeleiteten Raum*[8]) $\mathfrak{T}^*$, indem man zu Punkten aus $\mathfrak{T}^*$ die Mengen $\mathfrak{C}_\sigma$ erklärt, während offene Mengen in $\mathfrak{T}^*$ alle und nur die Mengen von Mengen $\mathfrak{C}_\sigma$ sind, deren Vereinigungsmenge in $\mathfrak{T}$ offen ist. Dann ist $\mathfrak{T}^*$ wieder ein topologischer Raum, da $\mathfrak{T}$ es ist und die Zerlegung $\{\mathfrak{C}_\sigma\}$ der Bedingung des Satzes 1. a) genügt.

Bildet man jetzt jeden Punkt $\mathfrak{p}$ auf die Kongruenzklasse $\mathfrak{C}_\pi$ ab, die $\mathfrak{p}$ enthält, so entsteht eine stetige Funktion α, und es ist $\mathfrak{T}^* = \alpha(\mathfrak{T})$ von minimaler Dichte bezgl. α.

ad b) Hierbei ist nur zu zeigen, daß $\mathfrak{T}^{**}$ wieder ein topologischer Raum ist, und dies folgt daraus, daß jede stetige Zerlegung die Bedingung von Satz 1. a) erfüllt.

ad c) Die Zuordnung

$$\alpha_1(\mathfrak{p}) \leftrightarrow \alpha_2(\mathfrak{p})$$

ist eineindeutig, da die Kongruenzklassen nach α_1 und α_2 dieselben sind. Ist nun $\mathfrak{M}_1$ irgendeine offene Menge aus $\alpha_1(\mathfrak{T})$, so ist ihr Urbild $\mathfrak{M}$ in $\mathfrak{T}$ offen, da α_1 stetig ist. Ist $\mathfrak{M}_2$ die $\mathfrak{M}_1$ vermöge obiger Zuordnung zugeordnete Teilmenge von $\alpha_2(\mathfrak{T})$, so ist $\mathfrak{M}_2 = \alpha_2(\mathfrak{M})$ und $\mathfrak{M}$ das Urbild von $\mathfrak{M}_2$. Da aber $\alpha_2(\mathfrak{T})$ von minimaler Dichte bezgl. α_2 ist und das Urbild $\mathfrak{M}$ von $\mathfrak{M}_2$ in $\mathfrak{T}$ offen ist, so ist auch $\mathfrak{M}_2$ in $\alpha_2(\mathfrak{T})$ offen, d. h. unsere Zuordnung bildet offene Mengen aus $\alpha_1(\mathfrak{T})$ auf offene Mengen aus $\alpha_2(\mathfrak{T})$ ab. Ebenso sieht man die Umkehrung ein, womit gezeigt ist, daß unsere Zuordnung eine topologische Abbildung von $\alpha_1(\mathfrak{T})$ auf $\alpha_2(\mathfrak{T})$ ist.

ad d) Jede Komponente einer stetigen Zerlegung ist abgeschlossen, da sie das Urbild einer abgeschlossenen Menge, nämlich eines Punktes ist.

---

[8]) Ähnlich wie bei Alexandroff a. a. O., S. 556—557, Definition 1.

Eine stetige Funktion α heißt *umkehrbar stetig*, wenn jeder offenen Menge aus $\mathfrak{T}$ eine offene Menge aus $\alpha(\mathfrak{T})$ entspricht. Es braucht dann aber nicht notwendig das Bild jeder abgeschlossenen Menge aus $\mathfrak{T}$ in $\alpha(\mathfrak{T})$ abgeschlossen zu sein[9]). Ist diese Bedingung für eine stetige Funktion erfüllt, so nennt man die Funktion *doppelstetig*[10]).

Jede umkehrbar stetige und ebenso jede doppelstetige Funktion bildet ihren Definitionsraum auf einen Bildraum von minimaler Dichte ab. Jedoch liefert nicht jede stetige Zerlegung umkehrbar stetige Funktionen. Hierfür ist vielmehr notwendig und hinreichend, daß man aus jeder offenen Menge aus $\mathfrak{T}$ durch Hinzufügung aller kongruenten Punkte wieder eine offene Menge erhält. Entsprechendes gilt für doppelstetige Funktionen.

§ 2.

Diskontinuitätsbereiche.

Definition 1. *Ist α eine in $\mathfrak{T}$ stetige Funktion, so heißt die in $\mathfrak{T}$ offene Menge $\mathfrak{D}$ ein α-Diskontinuitätsbereich, wenn*

1. *α auf $\mathfrak{D}$ topologisch und*
2. *$\overline{\alpha(\mathfrak{D})} = \alpha(\mathfrak{T})$ ist.*

Satz 2. *Dann und nur dann gibt es zu jedem Punkt aus $\mathfrak{T}$ einen ihn enthaltenden α-Diskontinuitätsbereich, wenn α im kleinen topologisch ist, d. h. wenn es zu jedem Punkt von $\mathfrak{T}$ eine Umgebung gibt, auf der α topologisch ist.*

Beweis. A. Gibt es zu jedem Punkte $\mathfrak{p}$ aus $\mathfrak{T}$ einen ihn enthaltenden α-Diskontinuitätsbereich $\mathfrak{D}$, so ist $\mathfrak{D}$ eine Umgebung von $\mathfrak{p}$, auf der α topologisch ist, d. h. α ist im kleinen topologisch.

B. Es sei α im kleinen topologisch. Wir beweisen dann zunächst den folgenden

Hilfssatz. *Ist α im kleinen topologisch, so gibt es zu jedem Punkte $\mathfrak{p}$ wenigstens eine Umgebung $\mathfrak{P}$, so daß α auf $\mathfrak{P}$, aber auf keiner echten, offenen Obermenge von $\mathfrak{P}$ topologisch ist.*

---

[9]) Dies zeigt das folgende Beispiel: Es sei $\mathfrak{K}_1$ die Teilmenge der x, y-Ebene, die aus sämtlichen Punkten (x, y) besteht, welche entweder $(x+2)^2+y^2<1$ oder $x<-2$ und $(x+2)^2+y^2=1$ erfüllen. Weiter entstehe $\mathfrak{K}_2$ aus $\mathfrak{K}_1$ durch Spiegelung an der y-Achse, und es sei $\mathfrak{T}=\mathfrak{K}_1 \vee \mathfrak{K}_2$. Die Festsetzung, daß zwei verschiedene Punkte dann und nur dann kongruent sind, wenn sie durch eine Verschiebung in Richtung der x-Achse um ± 4 auseinander hervorgehen, ergibt nach Satz 1. a) eine stetige Zerlegung von $\mathfrak{T}$, und die hierdurch definierte stetige Abbildung auf den aus der Zerlegung abgeleiteten Raum ist umkehrbar stetig und sogar [in der Bezeichnungsweise von § 2] im kleinen topologisch, aber nicht doppelstetig.

[10]) Vgl. Alexandroff a. a. O., S. 559, § 7.

Beweis. Wir nehmen eine Wohlordnung der sämtlichen in $\mathfrak{T}$ offenen Mengen vor. Da α im kleinen topologisch ist, gibt es Umgebungen von $\mathfrak{p}$, auf denen α topologisch ist; es sei $\mathfrak{U}_0$ die in der Wohlordnung erste von ihnen.

Wir konstruieren jetzt durch transfinite Induktion eine aufsteigende Kette von Umgebungen $\mathfrak{U}_\nu$ von $\mathfrak{p}$, auf denen α topologisch ist: Es sei bereits $\mathfrak{U}_\nu$ für alle $0 \leqq \nu < \mu$ definiert, so daß $\mathfrak{U}_\varrho > \mathfrak{U}_\sigma$ für $\varrho > \sigma$ und α auf allen $\mathfrak{U}_\nu$ topologisch ist. Ist dann μ keine Limeszahl, so ist entweder $\mathfrak{U}_{\mu-1}$ bereits die gesuchte Menge $\mathfrak{P}$, oder es gibt $\mathfrak{U}_{\mu-1}$ echt umfassende, offene Mengen, auf denen α topologisch ist; sei $\mathfrak{U}_\mu$ die erste unter ihnen. — Ist aber μ Limeszahl, so sei $\mathfrak{U}_\mu$ die Vereinigung aller $\mathfrak{U}_\nu$ mit $\nu < \mu$. Dann ist $\mathfrak{U}_\mu$ als Vereinigung offener Mengen wieder offen und α notwendig in $\mathfrak{U}_\mu$ topologisch. — Dieses Verfahren muß nach höchstens σ Schritten abbrechen, wobei σ der Ordnungstypus der wohlgeordneten Menge aller in $\mathfrak{T}$ offenen Mengen ist, und führt dann zu der gesuchten Menge $\mathfrak{P}$.

Um jetzt den Beweis von Satz 2 zu Ende zu führen, genügt es nachzuweisen, daß die Menge $\mathfrak{P}$ des Hilfssatzes ein Diskontinuitätsbereich ist, daß also

$$\overline{\alpha(\mathfrak{P})} = \alpha(\mathfrak{T})$$

ist. Wäre dies nicht wahr, so gäbe es einen Punkt $\mathfrak{p}^*$ aus $\alpha(\mathfrak{T}) = \mathfrak{T}^*$, der nicht Hp von $\alpha(\mathfrak{P})$ ist. Es gäbe also eine Umgebung $\mathfrak{U}^*$ von $\mathfrak{p}^*$ in $\mathfrak{T}^*$, so daß $\alpha(\mathfrak{P}) \cap \mathfrak{U}^* = 0$ ist. Das Urbild $\mathfrak{U}$ von $\mathfrak{U}^*$ ist offen, und es gilt $\mathfrak{P} \cap \mathfrak{U} = 0$. Da α auf $\mathfrak{T}$ im kleinen topologisch ist, so ist α auf einer offenen, nicht leeren Menge $\mathfrak{P}_1 \leqq \mathfrak{U}$ topologisch und wegen $\alpha(\mathfrak{P}) \cap \alpha(\mathfrak{P}_1) = 0$ auch auf $\mathfrak{P} \vee \mathfrak{P}_1 > \mathfrak{P}$, im Widerspruch zu der im Hilfssatz ausgesprochenen Eigenschaft von $\mathfrak{P}$.

Satz 3. *Die in $\mathfrak{T}$ stetige Funktion α ist dann und nur dann im kleinen topologisch, wenn α umkehrbar stetig und im kleinen eineindeutig ist.*

Zum Beweis genügt es einzusehen, daß eine im kleinen topologische Abbildung auch umkehrbar stetig ist. Sei also etwa $\mathfrak{M}$ in $\mathfrak{T}$ offen. Ist dann $\mathfrak{p}^*$ irgendein Punkt aus $\alpha(\mathfrak{M})$, $\mathfrak{p}$ ein beliebiges seiner Urbilder aus $\mathfrak{M}$, so gibt es eine Umgebung $\mathfrak{U}$ von $\mathfrak{p}$, auf der α topologisch ist. Dann ist auch $\mathfrak{U} \cap \mathfrak{M}$ eine Umgebung von $\mathfrak{p}$, auf der α topologisch ist, und also $\alpha(\mathfrak{U} \cap \mathfrak{M})$ eine in $\alpha(\mathfrak{T})$ offene Teilmenge von $\alpha(\mathfrak{M})$; d. h. $\alpha(\mathfrak{M})$ ist offen, wenn $\mathfrak{M}$ offen ist.

§ 3.

Vielfachheit.

Definition 2. *Ist $\mathfrak{M} \leqq \mathfrak{T}$ und $\mathfrak{p}$ ein Punkt aus $\mathfrak{T}$, so ist die Vielfachheit*

$$(\alpha(\mathfrak{p}), \mathfrak{M})$$

die Mächtigkeit der Menge der zu $\mathfrak{p}$ nach α kongruenten Punkte aus $\mathfrak{M}$.

Ist k irgendeine [endliche oder unendliche] Mächtigkeit, so bedeutet

$$\mathfrak{M}_{k,\alpha}$$

die Menge aller Punkte $\mathfrak{p}$ aus $\mathfrak{M}$, die $(\alpha(\mathfrak{p}), \mathfrak{M}) \geqq k$ erfüllen. Schließlich sei

$$\mathfrak{M}^{k,\alpha} = \mathfrak{M} - \mathfrak{M}_{k,\alpha}.$$

Satz 4. *Ist α umkehrbar stetig und k endlich, so ist $\mathfrak{T}_{k,\alpha}$ offen.*

Dieser Satz ist äquivalent mit dem folgenden:

4a. *Ist α umkehrbar stetig, k endlich und $(\alpha(\mathfrak{p}), \mathfrak{T}) \geqq k$, so hat $\mathfrak{p}$ eine Umgebung $\mathfrak{V}$, deren Punkte $\mathfrak{q}$ alle eine Vielfachheit $(\alpha(\mathfrak{q}), \mathfrak{V}) \geqq k$ aufweisen.*

Aus Satz 4 folgt nämlich 4a, wenn man darin $\mathfrak{V} = \mathfrak{T}_{k,\alpha}$ setzt. Ist umgekehrt 4a richtig, so ist $\mathfrak{T}_{k,\alpha}$ die Vereinigungsmenge der $\mathfrak{V}$-Umgebungen aller Punkte $\mathfrak{p}$, deren Vielfachheit $(\alpha(\mathfrak{p}), \mathfrak{T}) \geqq k$ ist. Unter dieser Bedingung ist also $\mathfrak{T}_{k,\alpha}$ als Vereinigungsmenge offener Mengen selbst offen.

Beweis von 4a. Es seien $\mathfrak{p}_1, \ldots, \mathfrak{p}_k$ endlich viel verschiedene, kongruente Punkte, und jedem $\mathfrak{p}_i$ werde eine Umgebung $\mathfrak{U}_i$ $(i = 1, \ldots, k)$ so zugeordnet, daß alle $\mathfrak{U}_i$ paarweise punktfremd sind. Die Mengen $\alpha(\mathfrak{U}_i)$ sind dann sämtlich in $\mathfrak{T}^* = \alpha(\mathfrak{T})$ Umgebungen von $\mathfrak{p}^* = \alpha(\mathfrak{p}_1) = \ldots = \alpha(\mathfrak{p}_k)$. Dasselbe gilt dann auch von $\mathfrak{V}^* = \alpha(\mathfrak{U}_1) \cap \ldots \cap \alpha(\mathfrak{U}_k)$. Jeder Punkt von $\mathfrak{V}^*$ hat dann mindestens k verschiedene Urbilder. Das Urbild $\mathfrak{V}$ von $\mathfrak{V}^*$ hat also die verlangte Eigenschaft. Hiermit ist die obige Behauptung und damit auch Satz 4 bewiesen.

Ist α im kleinen topologisch, so kann man die $\mathfrak{U}_i$ derart wählen, daß α auf ihnen topologisch ist. Dann ist jede Menge $\mathfrak{V}_i = \mathfrak{V} \cap \mathfrak{U}_i$ eine Umgebung von $\mathfrak{p}_i$, ferner α auf $\mathfrak{V}_i$ topologisch und $\alpha(\mathfrak{V}_i) = \mathfrak{V}^*$. Man erhält so den für später wichtigen

Zusatz 1. *Es sei α im kleinen topologisch; sind dann $\mathfrak{p}_1, \ldots, \mathfrak{p}_n$ endlich viele, verschiedene, nach α kongruente Punkte, so gibt es ein System punktfremder Mengen $\mathfrak{V}_1, \ldots, \mathfrak{V}_n$, so daß $\mathfrak{V}_i$ Umgebung von $\mathfrak{p}_i$, α auf $\mathfrak{V}_i$ topologisch und $\alpha(\mathfrak{V}_i) = \alpha(\mathfrak{V}_l)$ ist.*

Durch die Beziehung

$$\mathfrak{q}_i \longleftrightarrow \mathfrak{q}_l, \quad \text{falls} \quad \alpha(\mathfrak{q}_i) = \alpha(\mathfrak{q}_l) \quad \text{ist,}$$

werden also die Mengen $\mathfrak{V}_i$ und $\mathfrak{V}_l$ topologisch aufeinander abgebildet.

Für endliches k folgt aus Satz 4, daß $\mathfrak{T}^{k,\alpha}$ abgeschlossen ist. Wendet man ferner den Satz 4 auf eine beliebige Untermenge $\mathfrak{M}$ von $\mathfrak{T}$ an, so ergibt sich, daß $\mathfrak{M}_{k,\alpha}$ in $\mathfrak{M}$ offen ist. Wenn überdies $\mathfrak{M}$ in $\mathfrak{T}$ offen ist, so folgt hieraus auch, daß $\mathfrak{M}_{k,\alpha}$ in $\mathfrak{T}$ offen ist. Entsprechend erkennt man, daß aus der Abgeschlossenheit von $\mathfrak{M}$ in $\mathfrak{T}$ auch die Abgeschlossenheit von $\mathfrak{M}^{k,\alpha}$ in $\mathfrak{T}$ folgt.

Für unendliches k gelten Satz 4 und die obigen Folgerungen nicht mehr, wie das folgende Beispiel zeigt:

$\mathfrak{T}$ bestehe aus Punkten $\mathfrak{p}_{\mu\nu}$, wobei $1 \leqq \mu \leqq \omega$ [11]) und $0 \leqq \nu < \mu$ ist. In $\mathfrak{T}$ sind alle und nur die Mengen offen, die mit einem Punkte $\mathfrak{p}_{\omega\nu}$ auch fast alle Punkte $\mathfrak{p}_{\varrho\nu}$ mit $\nu < \varrho < \omega$ enthalten. Dann ist die Abbildung $\alpha(\mathfrak{p}_{\mu\nu}) = \mu$ von $\mathfrak{T}$ auf die Reihe der Ordinalzahlen von 1 bis ω stetig und im kleinen topologisch. Dagegen besteht $\mathfrak{T}^{\aleph_0, \alpha}$ aus allen und nur den Punkten $\mathfrak{p}_{\mu\nu}$ mit endlichem μ, während jedes $\mathfrak{p}_{\omega\nu}$ Hp von $\mathfrak{T}^{\aleph_0, \alpha}$ ist. $\mathfrak{T}^{\aleph_0, \alpha}$ ist mithin nicht abgeschlossen, $\mathfrak{T}_{\aleph_0, \alpha}$ also nicht offen.

Zusatz 2. *Ist $\mathfrak{T}$ kompakt* [12]), *α im kleinen topologisch und wenigstens eine der beiden Voraussetzungen*

a) *$\mathfrak{T}$ ist ein L-Raum* [13])

b) *α ist doppelstetig* [13a])

erfüllt, so bildet für jedes endliche n die Gesamtheit $\mathfrak{T}_{n,\alpha} - \mathfrak{T}_{n+1,\alpha}$ aller Punkte $\mathfrak{p}$ der Vielfachheit $(\alpha(\mathfrak{p}), \mathfrak{T}) = n$ eine zugleich offene und abgeschlossene Menge.

Beweis. Da α im kleinen topologisch ist, so kann keine Kongruenzklasse nach α einen Hp haben, muß also wegen der Kompaktheit von $\mathfrak{T}$ endlich sein. Die Zahlen $(\alpha(\mathfrak{p}), \mathfrak{T})$ sind also sämtlich endlich. Ferner muß nach Satz 3 und 4 jedes $\mathfrak{T}_{n,\alpha}$ offen, jedes $\mathfrak{T}^{n,\alpha}$ abgeschlossen sein, und da

$$\mathfrak{T}_{n,\alpha} - \mathfrak{T}_{n+1,\alpha} = \mathfrak{T}_{n,\alpha} \cap \mathfrak{T}^{n+1,\alpha}$$

ist, so genügt es also zu zeigen, daß $\mathfrak{T}_{n,\alpha}$ abgeschlossen ist, oder was hierfür hinreichend ist, daß für $n \geqq 2$ kein Punkt aus $\mathfrak{T}_{n-1,\alpha} - \mathfrak{T}_{n,\alpha}$ Hp von $\mathfrak{T}_{n,\alpha}$ ist. [$\mathfrak{T}_{1,\alpha} = \mathfrak{T}$ ist natürlich stets abgeschlossen.]

Es sei $\mathfrak{p}_1, \ldots, \mathfrak{p}_{n-1}$ eine Klasse kongruenter Punkte aus $\mathfrak{T}_{n-1,\alpha} - \mathfrak{T}_{n,\alpha}$; dann gibt es, da α im kleinen topologisch ist, untereinander punktfremde Umgebungen $\mathfrak{U}_i$ von $\mathfrak{p}_i$, auf denen α topologisch ist.

ad a) Ist $\mathfrak{T}$ ein L-Raum und $\mathfrak{p}_1$ Hp von $\mathfrak{T}_{n,\alpha}$, so gibt es eine Punktfolge $\mathfrak{a}_j$ aus $\mathfrak{T}_{n,\alpha}$ mit $\lim \mathfrak{a}_j = \mathfrak{p}_1$. Notwendig gibt es dann eine Folge $\mathfrak{b}_j$, so daß $\mathfrak{b}_j$ zu $\mathfrak{a}_j$ kongruent und in keinem $\mathfrak{U}_i$ enthalten ist. Ist dann $\mathfrak{a}_j^* = \alpha(\mathfrak{a}_j)$, so gibt es, da $\alpha(\mathfrak{T}) = \mathfrak{T}^*$ mit $\mathfrak{T}$ ein kompakter L-Raum ist und die $\mathfrak{a}_j^*$ eine unendliche Menge bilden müssen [zu jedem $\mathfrak{a}_j$ gibt es nur endlich viel kongruente], eine konvergente Teilfolge $\mathfrak{c}_j^*$ der Menge $\{\mathfrak{a}_j^*\}$. Ist dann $\lim \mathfrak{c}_j^* = \mathfrak{a}^*$, so muß $\alpha(\mathfrak{p}_1) = \mathfrak{a}^*$ sein; die $\mathfrak{b}_j$ müssen also einen zu $\mathfrak{p}_1$ kongruenten Hp haben, der in keinem $\mathfrak{U}_i$ liegt; das ist aber unmöglich.

---

[11]) ω ist die Anfangszahl der zweiten Zahlklasse.

[12]) D. h. jede unendliche Teilmenge von $\mathfrak{T}$ hat wenigstens einen Hp in $\mathfrak{T}$.

[13]) Vgl. F. Hausdorff, Mengenlehre, II. Aufl., Berlin 1927, S. 230.

[13a]) Diese Voraussetzung ist z. B. erfüllt, wenn $\mathfrak{T}$ bikompakt ist; vgl. R. Baer, Das Verhalten der Dimension topologischer Räume bei umkehrbar stetigen Abbildungen, Satz 2; erscheint im Journ. f. d. r. u. angew. Math.

ad b) Es sei α doppelstetig und $\mathfrak{p}_1$ Hp von $\mathfrak{T}_{n,\alpha}$. Zu jedem Punkte aus $\mathfrak{T}_{n,\alpha}$, der in einem $\mathfrak{U}_i$ vorkommt, gibt es in

$$\mathfrak{S} = \mathfrak{T}_{n,\alpha} - [\mathfrak{T}_{n,\alpha} \cap (\mathfrak{U}_1 \vee \ldots \vee \mathfrak{U}_{n-1})]$$

einen kongruenten. Dagegen enthält $\overline{\mathfrak{S}}$ die Punkte $\mathfrak{p}_i$ nicht. Da aber das Bild von $\overline{\mathfrak{S}}$ wegen der Doppelstetigkeit von α genau die abgeschlossene Hülle des Bildes von $\mathfrak{T}_{n,\alpha}$ ist:

$$\alpha(\overline{\mathfrak{S}}) = \overline{\alpha(\mathfrak{T}_{n,\alpha})},$$

so müßte $\alpha(\mathfrak{p}_1)$ auch in $\alpha(\overline{\mathfrak{S}})$ liegen, was unmöglich ist.

Daß ohne die Voraussetzungen a) *oder* b) *der Zusatz* 2 *im allgemeinen nicht richtig ist, zeigt das folgende Beispiel*:

$\mathfrak{T}$ bestehe aus Punkten $\mathfrak{p}_{i\nu}$ mit $i = 1, 2$ und $1 \leqq \nu \leqq \Omega$ [14]) für $i = 1$, $1 \leqq \nu < \Omega$ für $i = 2$. Weiter sei $\lim \mathfrak{p}_{i\nu} = \mathfrak{p}_{i\sigma}$, wenn $\lim \nu = \sigma$ ist. Dann ist $\mathfrak{T}$ ein kompakter Raum.

Ist weiter α die durch $\alpha(\mathfrak{p}_{i\nu}) = \nu$ definierte Abbildung von $\mathfrak{T}$ auf die Reihe der Ordinalzahlen $1 \leqq \nu \leqq \Omega$, so ist α im kleinen topologisch, und es besteht $\mathfrak{T}_{1,\alpha} - \mathfrak{T}_{2,\alpha}$ allein aus dem Punkt $\mathfrak{p}_{1,\Omega}$, der Hp der Menge $\mathfrak{T}_{2,\alpha}$ aller übrigen Punkte ist.

Zusatz 3. *Ist $\mathfrak{T}$ kompakt, α im kleinen topologisch und für jedes endliche n die Menge $\mathfrak{T}_{n,\alpha} - \mathfrak{T}_{n+1,\alpha}$ offen* [*und abgeschlossen*], *so ist die Reihe der in $\mathfrak{T}$ auftretenden Vielfachheiten nach oben durch eine endliche Zahl beschränkt.*

Beweis. Wie schon beim Beweis vom Zusatz 2 gezeigt wurde, ist $(\alpha(\mathfrak{p}), \mathfrak{T})$ stets endlich. Wäre nun unsere Behauptung unwahr, so gäbe es zu jeder endlichen Zahl n wenigstens einen Punkt $\mathfrak{p}_n$ mit $(\alpha(\mathfrak{p}_n), \mathfrak{T}) \geqq n$. Wegen der Kompaktheit von $\mathfrak{T}$ hat eine solche unendliche Menge $\{\mathfrak{p}_n\}$ wenigstens einen Hp $\mathfrak{q}$. Ist $(\alpha(\mathfrak{q}), \mathfrak{T}) = m$, so können in der offenen Menge $\mathfrak{T}_{m,\alpha} - \mathfrak{T}_{m+1,\alpha}$ nur Punkte $\mathfrak{p}_n$ mit $n \leqq m$, also nur endlich viele liegen, im Widerspruch zu der Annahme, daß $\mathfrak{q}$ ein Hp der Menge $\{\mathfrak{p}_n\}$ ist.

§ 4.

Blätterzahl und Überdeckungszahl.

Definition 3. *Ist $\mathfrak{M} \leqq \mathfrak{T}$ und α eine Funktion in $\mathfrak{T}$, so heiße die kleinste Mächtigkeit, die für jedes $\mathfrak{p}$ aus $\mathfrak{M}$ nicht kleiner als die Vielfachheit $(\alpha(\mathfrak{p}), \mathfrak{M})$ ist, die Blätterzahl*

$$B(\alpha, \mathfrak{M})$$

von α bezgl. $\mathfrak{M}$.

[14]) Ω ist die Anfangszahl der dritten Zahlklasse.

Die Blätterzahl braucht selbst nicht als Vielfachheit in $\mathfrak{M}$ aufzutreten [z. B. wenn die Vielfachheiten alle endlich, aber nicht beschränkt sind][15]).

Diese Blätterzahl steht in einem engen Zusammenhang mit der Anzahl der α-Diskontinuitätsbereiche, mit denen man $\mathfrak{M}$ überdecken kann.

Es seien

$$\mathfrak{U}_1, \mathfrak{U}_2, \ldots, \mathfrak{U}_\sigma, \ldots$$

punktfremde, offene Mengen in endlicher oder unendlicher Anzahl k [16]), auf denen α topologisch ist, und es sei

$$\mathfrak{M} \leqq \overline{\mathfrak{U}_1 \vee \ldots \vee \mathfrak{U}_\sigma \vee \ldots};$$

dann bilden die $\mathfrak{U}_\nu$ eine *α-Überdeckung* von $\mathfrak{M}$.

Gibt es überhaupt α-Überdeckungen von $\mathfrak{M}$, so auch solche, für welche die Mächtigkeit k einen kleinsten Wert annimmt. Diese Mächtigkeit heiße die *Überdeckungszahl*

$$U(\alpha, \mathfrak{M}).$$

Satz 5. *Ist α im kleinen topologisch, so gilt:*

a) *Es existiert die Überdeckungszahl $U(\alpha, \mathfrak{T})$.*

b) *Ist von den Zahlen $U(\alpha, \mathfrak{T})$ und $B(\alpha, \mathfrak{T})$ wenigstens eine endlich, so ist*

$$U(\alpha, \mathfrak{T}) = B(\alpha, \mathfrak{T}).$$

c) *Ist α doppelstetig, $U(\alpha, \mathfrak{T}) = \aleph_\nu$ und $B(\alpha, \mathfrak{T}) = \aleph_\beta$, so gilt*

$$\aleph_\nu \leqq \aleph_{\beta+1}.$$

[Ist also insbesondere $\nu = 0$ oder eine Limeszahl, so gilt $U(\alpha, \mathfrak{T}) \leqq B(\alpha, \mathfrak{T})$.]

Beweis. ad a) Wir nehmen eine Wohlordnung der sämtlichen offenen Teilmengen von $\mathfrak{T}$ vor und konstruieren dann durch transfinite Induktion eine α-Überdeckung von $\mathfrak{T}$.

Nach Satz 2 gibt es α-Diskontinuitätsbereiche. Es sei $\mathfrak{U}_1$ in der wohlgeordneten Reihe der offenen Mengen der erste α-Diskontinuitätsbereich.

Es sei $\mathfrak{U}_\nu$ für alle $1 \leqq \nu < \mu$ bereits konstruiert,

$$\mathfrak{M}_\mu = \mathfrak{T} - \overline{\mathfrak{U}_1 \vee \ldots \vee \mathfrak{U}_\nu \vee \ldots} \quad \text{für} \quad 1 \leqq \nu < \mu.$$

Dann gibt es in dem Unterraum $\mathfrak{M}_\mu$ von $\mathfrak{T}$, falls $\mathfrak{M}_\mu \neq 0$ ist, nach Satz 2 α-Diskontinuitätsbereiche, die in $\mathfrak{T}$ offen sind, da $\mathfrak{M}_\mu$ offen ist. Es sei $\mathfrak{U}_\mu$ in der wohlgeordneten Reihe der offenen Mengen der erste α-Diskontinuitätsbereich von $\mathfrak{M}_\mu$.

Dieses Verfahren wird fortgesetzt, bis einmal $\mathfrak{M}_\mu = 0$ wird, was für einen Index μ eintreten muß, der höchstens gleich dem Ordnungstypus der wohlgeordneten Reihe der offenen Mengen ist.

---

[15]) In dieser Sprechweise besagt der Zusatz 3 des § 3, daß $B(\alpha, \mathfrak{T})$ endlich ist.

[16]) k ist also die Mächtigkeit der Mengenmenge $\{\mathfrak{U}_\sigma\}$.

Nach Konstruktion bilden dann diese Mengen $\mathfrak{U}_\nu$ eine α-Überdeckung von $\mathfrak{T}$.

ad b) 1. *Es sei* $B(\alpha, \mathfrak{T})$ *endlich. Dann gilt*:

$$B(\alpha, \mathfrak{T}) \geqq U(\alpha, \mathfrak{T}).$$

Dies wird durch vollständige Induktion nach der endlichen Zahl $B(\alpha, \mathfrak{T}) = n$ gezeigt.

Für $n = 1$ gilt obige Ungleichung, da dann $\mathfrak{T}$ selbst eine α-Überdeckung von $\mathfrak{T}$ ist.

Es sei also die Ungleichung für $B(\alpha, \mathfrak{T}) = n - 1$ bereits bewiesen. Wir haben sie daraus für $B(\alpha, \mathfrak{T}) = n$ herzuleiten.

Nach Satz 2 gibt es in $\mathfrak{T}$ einen α-Diskontinuitätsbereich $\mathfrak{U}$. Dann ist $\mathfrak{S} = \mathfrak{T} - \overline{\mathfrak{U}}$ offen. Unsere Ungleichung ist bewiesen, wenn $U(\alpha, \mathfrak{S}) \leqq n - 1$ ist. Nach der Induktionsvoraussetzung genügt es also, $B(\alpha, \mathfrak{S}) \leqq n - 1$ oder $\mathfrak{S}_{n,\alpha} = 0$ zu beweisen.

Nach Satz 4 ist $\mathfrak{S}_{n,\alpha}$ und also auch $\alpha(\mathfrak{S}_{n,\alpha})$ offen. Andererseits kann kein Punkt aus $\mathfrak{U}$ zu Punkten aus $\mathfrak{S}_{n,\alpha}$ kongruent sein, da sonst $B(\alpha, \mathfrak{T}) \geqq n + 1$ wäre. Also ist

$$\alpha(\mathfrak{S}_{n,\alpha}) \leqq \mathfrak{T}^* - \alpha(\mathfrak{U}) = \overline{\alpha(\mathfrak{U})} - \alpha(\mathfrak{U}),$$

d. h. $\alpha(\mathfrak{S}_{n,\alpha})$ ist eine offene Menge ohne innere Punkte, also leer. Folglich muß auch $\mathfrak{S}_{n,\alpha} = 0$ sein, und unsere Ungleichung ist bewiesen[17]).

2. *Es sei* $U(\alpha, \mathfrak{T})$ *endlich. Dann gilt*:

$$U(\alpha, \mathfrak{T}) \geqq B(\alpha, \mathfrak{T}).$$

Dies wird durch vollständige Induktion nach der endlichen Zahl $U(\alpha, \mathfrak{T}) = m$ gezeigt.

Ist zunächst $U(\alpha, \mathfrak{T}) = 1$, so gibt es eine offene Menge $\mathfrak{U}$, auf der α topologisch ist, so daß $\overline{\mathfrak{U}} = \mathfrak{T}$ ist. Dann muß $B(\alpha, \mathfrak{T}) = 1$ sein. Ist nämlich $\mathfrak{p}_1$ ein Hp von $\mathfrak{U}$, der nicht zu $\mathfrak{U}$ gehört, und $\mathfrak{p}_2$ ein Punkt aus $\mathfrak{U}$, so kann $\mathfrak{p}_1$ nicht zu $\mathfrak{p}_2$ kongruent sein. Denn sonst gäbe es nach Zusatz 1 zu Satz 4 punktfremde Umgebungen $\mathfrak{U}_i$ von $\mathfrak{p}_i$, auf denen α topologisch ist, so daß $\alpha(\mathfrak{U}_1) = \alpha(\mathfrak{U}_2)$ und $\mathfrak{U}_2 \leqq \mathfrak{U}$ ist. Dies ist aber unmöglich, da $\mathfrak{U}_1 \cap \mathfrak{U} \neq 0$ ist. Also ist kein Punkt aus $\mathfrak{T} - \mathfrak{U}$ zu einem Punkt aus $\mathfrak{U}$ kongruent und also $\mathfrak{T}_{2,\alpha} \leqq \mathfrak{T} - \mathfrak{U}$. Da aber $\mathfrak{T}_{2,\alpha}$ nach Satz 4 offen ist und $\overline{\mathfrak{U}} = \mathfrak{T}$ gilt, so muß $\mathfrak{T}_{2,\alpha} = 0$, d. h. $B(\alpha, \mathfrak{T}) = 1$ sein.

Es sei also die Ungleichung für $U(\alpha, \mathfrak{T}) = m - 1$ bereits bewiesen; wir haben sie für $U(\alpha, \mathfrak{T}) = m$ daraus herzuleiten.

---

[17]) Man kann auch leicht die Existenz einer α-Überdeckung $\mathfrak{U}_1, \ldots, \mathfrak{U}_{B(\alpha,\mathfrak{T})}$ von $\mathfrak{T}$ derart einsehen, daß $\mathfrak{U}_1$ α-Diskontinuitätsbereich, $\alpha(\mathfrak{U}_i) \leqq \alpha(\mathfrak{U}_1)$ ist, und daß $\mathfrak{U}_i$ zu jedem Punkte $\mathfrak{q}$ aus $\mathfrak{U}_1$ mit $(\alpha(\mathfrak{q}), \mathfrak{T}) \geqq i$ wenigstens einen kongruenten enthält.

Es sei $\mathfrak{U}_1, \ldots, \mathfrak{U}_m$ eine α-Überdeckung von $\mathfrak{T}$, und $\mathfrak{A} = \overline{\mathfrak{U}_1 \vee \ldots \vee \mathfrak{U}_{m-1}}$. Dann ist $U(\alpha, \mathfrak{A}) \leqq m-1$, also nach der Induktionsannahme $B(\alpha, \mathfrak{A}) \leqq U(\alpha, \mathfrak{A})$. Weiter ist nach dem für $m=1$ bewiesenen $B(\alpha, \overline{\mathfrak{U}_m}) = U(\alpha, \overline{\mathfrak{U}_m}) = 1$. Hieraus folgt wegen

$$\mathfrak{T} = \overline{\mathfrak{U}_1 \vee \ldots \vee \mathfrak{U}_{m-1} \vee \mathfrak{U}_m} = \overline{\mathfrak{A} \vee \mathfrak{U}_m} = \overline{\mathfrak{A}} \vee \overline{\mathfrak{U}_m} = \mathfrak{A} \vee \overline{\mathfrak{U}_m},$$

daß $B(\alpha, \mathfrak{T}) \leqq U(\alpha, \mathfrak{A}) + 1 \leqq U(\alpha, \mathfrak{T})$ ist.

Es ist also gezeigt worden:

Ist $U(\alpha, \mathfrak{T})$ endlich, so ist $U(\alpha, \mathfrak{T}) \geqq B(\alpha, \mathfrak{T})$ und ist $B(\alpha, \mathfrak{T})$ endlich, so ist $B(\alpha, \mathfrak{T}) \geqq U(\alpha, \mathfrak{T})$. Ist also eine der beiden Zahlen endlich, so sind sie gleich.

ad c) Es sei α doppelstetig. Wir zeigen dann, daß die ad a) konstruierte α-Überdeckung von $\mathfrak{T}$ eine Mächtigkeit besitzt, die höchstens $\aleph_{\beta+1}$ ist.

Hierzu wird zunächst gezeigt, daß jeder Punkt aus $\mathfrak{U}_\varrho$ für $\sigma < \varrho$ einem Punkte aus $\mathfrak{M}_\sigma \wedge \overline{\mathfrak{U}}_\sigma$ kongruent ist. $\mathfrak{U}_\sigma$ war nämlich als α-Diskontinuitätsbereich des Unterraumes $\mathfrak{M}_\sigma$ ausgewählt worden. Also ist $\alpha(\mathfrak{M}_\sigma) = \overline{\alpha(\mathfrak{U}_\sigma)} \wedge \alpha(\mathfrak{M}_\sigma)$. Wegen der Doppelstetigkeit von α ist aber $\alpha(\overline{\mathfrak{U}_\sigma}) = \overline{\alpha(\mathfrak{U}_\sigma)}$ und sogar $\alpha(\mathfrak{M}_\sigma) = \alpha(\overline{\mathfrak{U}}_\sigma \wedge \mathfrak{M}_\sigma)$. Daher ist jeder Punkt von $\mathfrak{M}_\sigma$, insbesondere also jeder Punkt von $\mathfrak{U}_\varrho \leqq \mathfrak{M}_\varrho \leqq \mathfrak{M}_\sigma$ zu einem Punkt aus $\mathfrak{M}_\sigma \wedge \overline{\mathfrak{U}}_\sigma$ kongruent.

Da aber die Mengen $\overline{\mathfrak{U}}_\sigma \wedge \mathfrak{M}_\sigma$ untereinander punktfremd sind, so gilt, wenn man unter $|\nu|$ die Mächtigkeit der Ordinalzahl ν versteht, für jeden Punkt $\mathfrak{p}$ aus $\mathfrak{U}_\nu$, daß $(\alpha(\mathfrak{p}), \mathfrak{T}) \geqq |\nu|$ ist. Ist also λ der Ordnungstypus der ad a) konstruierten α-Überdeckung von $\mathfrak{T}$, so gilt für jedes $\nu < \lambda$, daß $\aleph_\beta \geqq |\nu|$ ist. Also ist sicher $\aleph_{\beta+1} \geqq |\lambda|$, was zu beweisen war.

Folgendes Beispiel zeigt, daß Blätter- und Überdeckungszahl nicht mehr gleich zu sein brauchen, wenn sie unendlich sind:

$\mathfrak{T}$ bestehe aus Punkten $\mathfrak{p}_{\mu,\nu}$ mit $1 \leqq \mu \leqq \Omega$ und $0 \leqq \nu < \mu$. Eine Teilmenge $\mathfrak{M}$ von $\mathfrak{T}$ heißt offen, wenn es zu jedem Punkte $\mathfrak{p}_{\mu,\nu}$ aus $\mathfrak{M}$, wo μ Limeszahl ist, ein $\mu_0 < \mu$ gibt, so daß alle $\mathfrak{p}_{\varrho,\nu}$ mit $\mu_0 \leqq \varrho \leqq \mu$ zu $\mathfrak{M}$ gehören. Dann ist $\mathfrak{T}$ ein topologischer Raum, in dem insbesondere jeder Punkt $\mathfrak{p}_{\mu+1,\nu}$ eine offene Menge bildet.

Ist $\alpha(\mathfrak{p}_{\mu,\nu}) = \mu$, so ist α eine im kleinen topologische, doppelstetige Abbildung von $\mathfrak{T}$ auf die Reihe der Ordinalzahlen $1 \leqq \mu \leqq \Omega$. Dann gilt offenbar $B(\alpha, \mathfrak{T}) = \aleph_1$. Also muß auch $U(\alpha, \mathfrak{T})$ unendlich sein. Es bleibt also nur noch zu zeigen, daß $U(\alpha, \mathfrak{T}) \leqq \aleph_0$ ist.

Hierzu betrachten wir die Menge $\mathfrak{U}$ aller Punkte $\mathfrak{p}_{\mu,\nu}$, wo μ keine Limeszahl ist. Dann ist $\mathfrak{U}$ offen und $\overline{\mathfrak{U}} = \mathfrak{T}$. Es genügt also, eine abzählbare α-Überdeckung von $\mathfrak{U}$ anzugeben. Da für jeden Punkt $\mathfrak{p}_{\mu,\nu}$ aus $\mathfrak{U}$ gilt: $(\alpha(\mathfrak{p}_{\mu,\nu}), \mathfrak{T}) \leqq \aleph_0$, so ist es möglich, die zu $\mathfrak{p}_{\mu,\nu}$ nach α kongruenten

Punkte in eine [endliche oder unendliche] Reihe: $\mathfrak{p}_{\mu,\nu_1}, \mathfrak{p}_{\mu,\nu_2}, \mathfrak{p}_{\mu,\nu_3}, \ldots$ anzuordnen. Besteht dann $\mathfrak{U}_i$ für jede endliche Zahl $i \geqq 1$ aus den sämtlichen Punkten $\mathfrak{p}_{\mu,\nu_i}$ aus $\mathfrak{U}$, so ist $\mathfrak{U}_i$ offen, α auf $\mathfrak{U}_i$ topologisch und $\mathfrak{U} = \mathfrak{U}_1 \vee \ldots \vee \mathfrak{U}_i \vee \ldots$; die Mengen $\mathfrak{U}_i$ bilden also eine abzählbare α-Überdeckung von $\mathfrak{T}$, und es ist mithin:

$$U(\alpha, \mathfrak{T}) < B(\alpha, \mathfrak{T}).$$

Ist $\mathfrak{M}$ eine offene Teilmenge von $\mathfrak{T}$, so ist jede in $\mathfrak{M}$ offene Menge auch in $\mathfrak{T}$ offen. Wendet man also Satz 5 auf den Unterraum $\mathfrak{M}$ von $\mathfrak{T}$ für in $\mathfrak{T}$ offenes $\mathfrak{M}$ an, so erhält man entsprechende Aussagen über $U(\alpha, \mathfrak{M})$ und $B(\alpha, \mathfrak{M})$.

Will man sich von der Voraussetzung, daß α im kleinen topologisch ist, befreien, so wird man zwei Arten von Punkten aus $\mathfrak{T}$ zu unterscheiden haben: die *regulären*, das sind solche, zu denen es eine Umgebung gibt, auf der α topologisch ist, und die *singulären*, das sind solche, die nicht regulär sind. Offenbar ist die Gesamtheit $\mathfrak{R}$ der regulären Punkte eine offene Menge, die dann und nur dann mit $\mathfrak{T}$ zusammenfällt, wenn α im kleinen topologisch ist.

Dann und nur dann existiert die Überdeckungszahl $U(\alpha, \mathfrak{T})$, *wenn*

$$\overline{\mathfrak{R}} = \mathfrak{T}$$

ist.

Beweis. Existiert einerseits die Überdeckungszahl, so gibt es wenigstens eine α-Überdeckung $\{\mathfrak{U}_\nu\}$ von $\mathfrak{T}$, und es ist $\mathfrak{U}_1 \vee \ldots \vee \mathfrak{U}_\nu \vee \ldots \leqq \mathfrak{R}$, also wegen $\overline{\mathfrak{U}_1 \vee \ldots \vee \mathfrak{U}_\nu \vee \ldots} = \mathfrak{T}$ auch $\overline{\mathfrak{R}} = \mathfrak{T}$.

Es sei umgekehrt $\overline{\mathfrak{R}} = \mathfrak{T}$. Dann ist α auf $\mathfrak{R}$ im kleinen topologisch, und es gibt nach Satz 5 a) eine α-Überdeckung von $\mathfrak{R}$. Da aber $\mathfrak{R}$ in $\mathfrak{T}$ offen und $\overline{\mathfrak{R}} = \mathfrak{T}$ ist, so ist die α-Überdeckung von $\mathfrak{R}$ auch eine von $\mathfrak{T}$, und es existiert $U(\alpha, \mathfrak{T})$.

Von den übrigen Aussagen des Satzes 5 läßt sich nur noch ein Teil übertragen: da α auf $\mathfrak{R}$ im kleinen topologisch ist und $\overline{\mathfrak{R}} = \mathfrak{T}$, so ist $U(\alpha, \mathfrak{T}) = U(\alpha, \mathfrak{R})$, dagegen ist $B(\alpha, \mathfrak{T}) \geqq B(\alpha, \mathfrak{R})$, und hier wird im allgemeinen das Gleichheitszeichen nicht gelten.

§ 5.

Zulässige Funktionen.

In dem Raume $\mathfrak{T}$ sei eine Gruppe T topologischer Abbildungen von $\mathfrak{T}$ auf sich ausgezeichnet. Eine in $\mathfrak{T}$ stetige Funktion α heiße [bezgl. T] *zulässig*, wenn für jedes τ aus T durch

$$\tau^*(\alpha(\mathfrak{p})) = \alpha(\tau(\mathfrak{p}))$$

eine topologische Abbildung τ^* von $\mathfrak{T}^* = \alpha(\mathfrak{T})$ auf sich definiert wird.

Hierfür gilt das folgende einfache Kriterium:

Ist $\mathfrak{T}^*$ bezgl. α von minimaler Dichte[18]*), so ist α dann und nur dann eine [bezgl. T] zulässige Funktion, wenn mit $\alpha(\mathfrak{p}_1) = \alpha(\mathfrak{p}_2)$ auch stets $\alpha(\tau(\mathfrak{p}_1)) = \alpha(\tau(\mathfrak{p}_2))$ gilt.*

Beweis. A. Erfüllt α die obige Bedingung, ist $\mathfrak{T}^*$ von minimaler Dichte bezgl. α und $\mathfrak{p}^*$ irgendein Punkt aus $\mathfrak{T}^*$, so gibt es wenigstens einen Punkt $\mathfrak{p}$ aus $\mathfrak{T}$ mit $\alpha(\mathfrak{p}) = \mathfrak{p}^*$. Dann ist

$$\tau^*(\mathfrak{p}^*) = \tau^*(\alpha(\mathfrak{p})) = \alpha(\tau(\mathfrak{p}))$$

ein Bildpunkt von $\mathfrak{p}^*$ bei τ^*, d. h. τ^* ist in ganz $\mathfrak{T}^*$ definiert.

Ist weiter $\mathfrak{q}$ ein zweiter Punkt aus $\mathfrak{T}$, so daß $\alpha(\mathfrak{q}) = \mathfrak{p}^*$ ist, so folgt aus der Erfülltheit von $\alpha(\tau(\mathfrak{q})) = \alpha(\tau(\mathfrak{p}))$ auch, daß $\tau^*(\alpha(\mathfrak{p})) = \tau^*(\alpha(\mathfrak{q}))$, d. h. τ^* eindeutig ist.

Ist endlich $\mathfrak{M}^*$ eine in $\mathfrak{T}^*$ offene Menge und $\mathfrak{M}$ das Urbild von $\mathfrak{M}^*$ in $\mathfrak{T}$ bei α, so ist wegen der Stetigkeit von α auch $\mathfrak{M}$ eine offene Menge. Betrachtet man jetzt $\tau^*(\mathfrak{M}^*)$, so folgt aus

$$\tau^*(\mathfrak{M}^*) = \tau^*(\alpha(\mathfrak{M})) = \alpha(\tau(\mathfrak{M}))$$

und der Umkehrbarkeit von τ, daß $\tau(\mathfrak{M})$ das Urbild von $\tau^*(\mathfrak{M}^*)$ ist. Da mit $\mathfrak{M}$ auch $\tau(\mathfrak{M})$ in $\mathfrak{T}$ offen und $\mathfrak{T}^*$ von minimaler Dichte in bezug auf α ist, so muß auch $\alpha(\tau(\mathfrak{M})) = \tau^*(\mathfrak{M}^*)$ in $\mathfrak{T}^*$ offen sein, d. h. τ^* bildet offene Mengen auf offene Mengen ab.

Da schließlich τ^{-1} in $\mathfrak{T}^*$ genau $(\tau^*)^{-1}$ induziert und für $(\tau^*)^{-1}$ dasselbe wie für τ^* gilt, so ist τ^* umkehrbar eindeutig und umkehrbar stetig, d. h. topologisch und α also zulässig.

B. Erfüllt α nicht die obige Bedingung, so gibt es ein τ aus T, so daß für ein geeignetes Punktepaar $\mathfrak{p}$, $\mathfrak{q}$ aus $\mathfrak{T}$ zwar $\alpha(\mathfrak{p}) = \alpha(\mathfrak{q})$, aber $\alpha(\tau(\mathfrak{p})) \neq \alpha(\tau(\mathfrak{q}))$ gilt. Dann wird einerseits

$$\tau^*(\mathfrak{p}^*) = \alpha(\tau(\mathfrak{p})),$$

andererseits

$$\tau^*(\mathfrak{p}^*) = \alpha(\tau(\mathfrak{q})),$$

d. h. τ^* ist mehrdeutig, α nicht zulässig.

Ist α eine bezgl. T zulässige Funktion, so ist die Gesamtheit $\mathfrak{D}$ aller in α-Diskontinuitätsbereichen enthaltenen Punkte aus $\mathfrak{T}$ eine offene, bezgl. T invariante Menge. Ist überdies α umkehrbar stetig und $\mathfrak{D} \neq 0$, so ist $\alpha(\mathfrak{T}) - \alpha(\mathfrak{D})$ ein Randstück[19]*) bezgl. $\alpha(\mathsf{T})$.*

Beweis. Ist τ eine beliebige Abbildung aus T und $\mathfrak{M}$ ein α-Diskontinuitätsbereich, so ist α als zulässige Abbildung auch auf $\tau(\mathfrak{M})$

[18]) Diese Voraussetzung ist natürlich nicht notwendig; doch wäre der Satz im allgemeinen ohne sie falsch.

[19]) Im Sinne von R. Baer - F. Levi a. a. O., S. 173, Definition 1.

topologisch, also $\tau(\mathfrak{M})$ mit $\mathfrak{M}$ α-Diskontinuitätsbereich, d. h. $\mathfrak{D}$ ist invariant gegenüber T. Ist weiter $\mathsf{T}^* = \alpha(\mathsf{T})$ und $\mathfrak{D}^* = \alpha(\mathfrak{D})$, so ist $\mathsf{T}^*(\mathfrak{D}^*) = \mathsf{T}^*(\alpha(\mathfrak{D})) = \alpha(\mathsf{T}(\mathfrak{D})) = \alpha(\mathfrak{D}) = \mathfrak{D}^*$, d. h. $\mathfrak{D}^*$ ist gegenüber T^* invariant.

Weiter ist $\mathfrak{D}$ als Vereinigung offener Mengen offen, also wegen der umkehrbaren Stetigkeit von α auch $\mathfrak{D}^*$.

Ist schließlich $\mathfrak{D} \neq 0$, so folgt aus der Definition der α-Diskontinuitätsbereiche, daß

$$\overline{\mathfrak{D}^*} = \overline{\alpha(\mathfrak{D})} = \alpha(\mathfrak{T}) = \mathfrak{T}^*$$

ist, mithin[20]) ist $\mathfrak{T}^* - \mathfrak{D}^*$ ein Randstück bezgl. T^*.

§ 6.

Automorphe Funktionen.

Im Raume $\mathfrak{T}$ sei eine Gruppe T topologischer Abbildungen von $\mathfrak{T}$ auf sich ausgezeichnet. *Dann heißt eine in $\mathfrak{T}$ stetige Funktion α automorph [bezgl. T], wenn für alle τ aus T gilt*

$$\alpha(\mathfrak{p}) = \alpha(\tau(\mathfrak{p})).^{21})$$

Satz 6. *Zu jedem Raume $\mathfrak{T}$ mit ausgezeichneter Gruppe T gibt es eine (im wesentlichen eindeutig bestimmte) automorphe Funktion φ, so daß*

$$\sigma(\varphi(\mathfrak{p}))$$

alle bezgl. T in $\mathfrak{T}$ automorphen Funktionen durchläuft, wenn σ alle in $\varphi(\mathfrak{T})$ stetigen Funktionen durchläuft[22]).

---

[20]) Nach R. Baer-F. Levi a. a. O., Satz 12, S. 181.

[21]) Automorphe Funktionen sind stets zulässige Funktionen. Für automorphe Funktionen α besteht die Gruppe $\alpha(\mathsf{T})$ allein aus der Identität.

[22]) Genau ebenso beweist man die folgende allgemeinere Aussage:

Ist Σ irgendeine Menge in $\mathfrak{T}$ stetiger Funktionen, so gibt es eine [im wesentlichen eindeutig bestimmte] in $\mathfrak{T}$ stetige Funktion φ [die nicht zu Σ zu gehören braucht], so daß

1. sich jede Funktion aus Σ in der Form

$$\sigma[\varphi(\mathfrak{p})]$$

darstellen läßt, wobei σ eine geeignete, in $\varphi(\mathfrak{T})$ stetige Funktion ist, und

2. für jede in $\mathfrak{T}$ stetige Funktion ψ derart, daß a) jede Funktion aus Σ sich in der Form $\sigma[\psi(\mathfrak{p})]$ mit in $\psi(\mathfrak{T})$ stetigem σ darstellen läßt und b) $\psi(\mathfrak{T})$ von minimaler Dichte ist, auch

$$\varphi(\mathfrak{p}) = \chi[\psi(\mathfrak{p})]$$

gilt, wobei χ eine in $\psi(\mathfrak{T})$ stetige Funktion ist.

Wählt man insbesondere als Menge Σ die Gesamtheit der zulässigen Funktionen, so kann man stets als 1. und 2. erfüllende Funktionen φ die identische Abbildung von $\mathfrak{T}$ auf sich wählen, während dies für die Gesamtheit der automorphen Funktionen im allgemeinen nicht möglich ist.

Beweis. Es gibt stets in $\mathfrak{T}$ automorphe Funktionen; z. B. ist die Abbildung aller Punkte von $\mathfrak{T}$ auf einen und denselben Punkt eine solche.

Ist $\mathfrak{p}$ irgendein Punkt aus $\mathfrak{T}$, so sei $\mathfrak{C}(\mathfrak{p})$ der Durchschnitt aller $\mathfrak{p}$ enthaltenden Komponenten von solchen stetigen Zerlegungen, die durch in $\mathfrak{T}$ bezgl. T automorphe Funktionen definiert werden.

Dann seien $\{\mathfrak{C}_\nu\}$ die verschiedenen unter den Mengen $\mathfrak{C}(\mathfrak{p})$.

Nach Konstruktion kommt dann jeder Punkt aus $\mathfrak{T}$ in wenigstens einer Menge $\mathfrak{C}_\nu$ vor. Weiter bilden die $\mathfrak{C}_\nu$ eine stetige Zerlegung von $\mathfrak{T}$. Sind nämlich $\mathfrak{C}_\nu$ und $\mathfrak{C}_\mu$ zwei verschiedene Komponenten, so gibt es, wie aus der Konstruktion der $\mathfrak{C}(\mathfrak{p})$ folgt, eine automorphe Funktion α, so daß die $\mathfrak{C}_\nu$ enthaltende Kongruenzklasse $\mathfrak{A}_\nu$ von der $\mathfrak{C}_\mu$ enthaltenden Kongruenzklasse $\mathfrak{A}_\mu$ verschieden ist. Da α als automorphe Funktion stetig ist, so gibt es nach Satz 1 punktfremde Umgebungen $\mathfrak{V}_\nu$ und $\mathfrak{V}_\mu$ von $\mathfrak{A}_\nu$ bzw. $\mathfrak{A}_\mu$, die aus vollen Kongruenzklassen nach α bestehen. Dann sind $\mathfrak{V}_\nu$ und $\mathfrak{V}_\mu$ auch punktfremde Umgebungen von $\mathfrak{C}_\nu$ bzw. $\mathfrak{C}_\mu$, die aus vollständigen Klassen $\mathfrak{C}_\varkappa$ bestehen. Nach Satz 1 des § 1 gibt es also eine stetige Funktion φ, deren Kongruenzklassen gerade die $\mathfrak{C}_\varkappa$ sind. φ ist auch eine automorphe Funktion, denn die $\mathfrak{C}_\varkappa$ sind als Durchschnitt bezgl. T invarianter Mengen selbst invariant.

Daß sich schließlich alle und nur die automorphen Funktionen auf die Form $\sigma(\varphi(\mathfrak{p}))$ mit in $\varphi(\mathfrak{T})$ stetigem σ bringen lassen, folgt daraus, daß die Komponenten der zu irgendeiner automorphen Funktion gehörenden stetigen Zerlegung durch Zusammenfassung von Kongruenzklassen nach φ gewonnen werden[23]).

---

[23]) In der Bezeichnungsweise von R. Baer - F. Levi, a. a. O., S. 174, Definition 2 gilt, daß mit jedem $\mathfrak{p}$ auch das Raumelement $E(\mathfrak{p})$ zu einer Kongruenzklasse nach φ gehört, da diese ja abgeschlossen und invariant sind. Ist daher $\mathfrak{p}$ in $E(\mathfrak{q})$ enthalten, so muß wegen der Punktfremdheit verschiedener Kongruenzklassen mit $\mathfrak{p}$ auch $E(\mathfrak{q})$ zur selben Kongruenzklasse gehören. Doch brauchen die Kongruenzklassen nicht die kleinsten, abgeschlossenen, invarianten Mengen mit dieser Eigenschaft zu sein, wie das folgende Beispiel zeigt:

$\mathfrak{T}$ sei die folgende Teilmenge der x, y-Ebene:

$$0 < x < 1 \quad \text{für} \quad y = \frac{1}{n}, \qquad x = 0 \text{ oder } = 1 \quad \text{für} \quad y = 0.$$

T bestehe aus allen topologischen Abbildungen von $\mathfrak{T}$ auf sich, bei denen die Richtung und der Abstand von der x-Achse erhalten bleibt. Kongruenzklassen nach φ bilden dann alle Punkte mit gleicher y-Koordinate. Dagegen hat die aus dem Punkte $(0, 0)$ allein bestehende Menge die Eigenschaft invariant, abgeschlossen zu sein und mit irgendeinem Punkte $\mathfrak{p}$ auch jedes $\mathfrak{p}$ enthaltende $E(\mathfrak{q})$ zu enthalten.

§ 7.

Diskontinuierliche Gruppen.

Definition 4. *Eine Gruppe* T *topologischer Abbildungen von* $\mathfrak{T}$ *auf sich heiße diskontinuierlich, wenn es auf* $\mathfrak{T}$ *eine stetige Funktion* α *gibt, so daß*

1. *die* α*-Kongruenzklassen mit den* T*-Äquivalenzklassen identisch sind und*

2. *jeder Punkt aus* $\mathfrak{T}$ *in einem* α*-Diskontinuitätsbereich enthalten ist.*

Satz 7. *Dann und nur dann ist* T *diskontinuierlich, wenn es*

a) *zu jedem Punkte eine Umgebung gibt, die kein Paar verschiedener äquivalenter Punkte enthält, und es*

b) *zu je zwei nicht äquivalenten Punkten* $\mathfrak{p}$ *und* $\mathfrak{q}$ [*punktfremde*] *Umgebungen* $\mathfrak{U}(\mathfrak{p})$ *und* $\mathfrak{U}(\mathfrak{q})$ *gibt, so daß kein Punkt von* $\mathfrak{U}(\mathfrak{p})$ *zu einem Punkte von* $\mathfrak{U}(\mathfrak{q})$ *äquivalent ist*[24])[25]).

Beweis. A. Es sei T diskontinuierlich. Dann ist die Zerlegung von $\mathfrak{T}$ in T-Äquivalenzklassen eine stetige Zerlegung und also wegen Satz 1 die Bedingung b) erfüllt. Da dann weiter nach Satz 2 die zugehörige Funktion α im kleinen topologisch ist, so ist sie a fortiori im kleinen eineindeutig und also auch die Bedingung a) erfüllt.

B. Es seien die Bedingungen a) und b) des Satzes erfüllt. Sind dann $\mathfrak{p}$ und $\mathfrak{q}$ zwei nicht äquivalente Punkte, so gibt es wegen b) Umgebungen $\mathfrak{U}(\mathfrak{p})$ und $\mathfrak{U}(\mathfrak{q})$, so daß kein Punkt von $\mathfrak{U}(\mathfrak{p})$ mit einem von $\mathfrak{U}(\mathfrak{q})$ äquivalent ist. Es seien $\mathfrak{C}(\mathfrak{p})$ bzw. $\mathfrak{C}(\mathfrak{q})$ die aus $\mathfrak{U}(\mathfrak{p})$ bzw. $\mathfrak{U}(\mathfrak{q})$ durch Hinzufügung aller äquivalenten Punkte entstehenden Mengen. Dann sind $\mathfrak{C}(\mathfrak{p})$ und $\mathfrak{C}(\mathfrak{q})$ offen, punktfremd und enthalten vollständige T-Äquivalenzklassen. Mithin ist die Zerlegung in T-Äquivalenzklassen nach Satz 1 eine stetige Zerlegung, und es sei α eine hierdurch definierte Funktion derart, daß $\alpha(\mathfrak{T})$ von minimaler Dichte in bezug auf α ist.

Wegen der Bedingung a) ist α im kleinen eineindeutig. Ist weiter $\mathfrak{M}$ irgendeine in $\mathfrak{T}$ offene Menge, so ist die Gesamtheit $\mathfrak{N}$ aller zu Punkten aus $\mathfrak{M}$ äquivalenten Punkte ebenfalls offen und das Urbild von $\alpha(\mathfrak{M})$. Also ist $\alpha(\mathfrak{M})$ offen, da ja $\alpha(\mathfrak{T})$ von minimaler Dichte in bezug auf α ist. Mithin ist α auch umkehrbar stetig.

---

[24]) Dieser Satz ordnet die geläufigen Sätze von der Existenz der Diskontinuitätsbereiche von Bewegungsgruppen in den hier gegebenen Zusammenhang ein.

[25]) Ist T diskontinuierlich, so kann wegen Bedingung a) das Raumelement $E(\mathfrak{p})$ keinen Hp haben; insbesondere besteht also $E(\mathfrak{p})$ aus der Gesamtheit der zu $\mathfrak{p}$ äquivalenten Punkte.

Nach Satz 3 ist α also im kleinen topologisch, und nach Satz 2 existiert mithin zu jedem Punkt aus $\mathfrak{T}$ ein ihn enthaltender α-Diskontinuitätsbereich.

Die Betrachtungen über diskontinuierliche Gruppen gewinnen ein besonderes Interesse, wenn $\mathfrak{T}$ kompakt[26]) ist. Es gilt nämlich

Satz 8. *Ist $\mathfrak{T}$ kompakt und im kleinen zusammenhängend*[27]), *T diskontinuierlich, so ist T endlich.*

Satz 9. *Ist $\mathfrak{T}$ kompakt, T diskontinuierlich*[28]), *so ist T dann und nur dann endlich, wenn es zu jedem Punkt aus $\mathfrak{T}$ eine der folgenden Bedingung* (S) *genügende Umgebung $\mathfrak{U}$ gibt:*

(S) *Bei jeder Abbildung aus T bleibt $\mathfrak{U}$ entweder punktweise invariant oder geht in eine zu $\mathfrak{U}$ punktfremde Menge über.*

Satz 8 ist in Satz 9 enthalten, wie folgende Überlegung zeigt:

Da T diskontinuierlich ist, hat nach Satz 7 a) die Menge der zu einem Punkte äquivalenten Punkte keinen Hp, ist also wegen der Kompaktheit von $\mathfrak{T}$ endlich.

Versteht man, wie stets im folgenden, unter α die durch T definierte stetige Funktion, so ist α im kleinen topologisch, und es gibt nach obigem zu jedem Punkte $\mathfrak{p}$ eine Umgebung $\mathfrak{V}$, in der α topologisch ist, derart daß kein von $\mathfrak{p}$ verschiedener, zu $\mathfrak{p}$ äquivalenter Punkt Hp von $\mathfrak{V}$ ist. Da $\mathfrak{T}$ im kleinen zusammenhängend ist, ist in $\mathfrak{V}$ eine zusammenhängende Umgebung $\mathfrak{U}$ von $\mathfrak{p}$ enthalten.

Ist dann τ irgendeine Abbildung aus T, so ist mit $\mathfrak{U}$ auch $\tau(\mathfrak{U})$ und also auch $\mathfrak{U} \cap \tau(\mathfrak{U})$ offen. Ist nun $\mathfrak{U} \cap \tau(\mathfrak{U}) \neq 0$, so bewirkt τ auf $\mathfrak{U} \cap \tau(\mathfrak{U})$ die identische Abbildung, da α auf $\mathfrak{V}$, also auch auf $\mathfrak{U}$ topologisch ist und es daher in $\mathfrak{U}$ kein Paar verschiedener, äquivalenter Punkte geben kann. Also erfahren durch τ auch alle Hp von $\mathfrak{U} \cap \tau(\mathfrak{U})$ die identische Abbildung, d. h.

$$\mathfrak{U} \cap \tau(\mathfrak{U}) = \mathfrak{U} \cap \overline{[\mathfrak{U} \cap \tau(\mathfrak{U})]},$$

d. h. $\mathfrak{U} \cap \tau(\mathfrak{U})$ ist in $\mathfrak{U}$ offen und abgeschlossen. Da $\mathfrak{U}$ zusammenhängend und $\mathfrak{U} \cap \tau(\mathfrak{U}) \neq 0$ ist, so muß also $\mathfrak{U} \cap \tau(\mathfrak{U}) = \mathfrak{U}$, d. h. $\mathfrak{U} = \tau(\mathfrak{U})$ sein; $\mathfrak{U}$ erfüllt also die Bedingung (S).

---

[26]) Vgl. Fußnote [12]).

[27]) Ein Raum heißt nach H. Hahn [vgl. F. Hausdorff, Mengenlehre, II. Aufl., Berlin 1927, S. 155] im kleinen zusammenhängend, wenn es zu jedem Punkte $\mathfrak{p}$ und jeder Umgebung $\mathfrak{U}$ von $\mathfrak{p}$ eine in $\mathfrak{U}$ enthaltene zusammenhängende Umgebung von $\mathfrak{p}$ gibt. Dabei heißt eine Menge $\mathfrak{M} \neq 0$ zusammenhängend, wenn sie sich nicht in zwei punktfremde, von 0 verschiedene, in $\mathfrak{M}$ abgeschlossene Teilmengen spalten läßt. Vgl. F. Hausdorff, a. a. O., S. 150.

[28]) Man bemerke, daß die Diskontinuität von T nicht aus der Kompaktheit von $\mathfrak{T}$ und Endlichkeit von T folgt, wie einfachste Beispiele zeigen.

Beweis von Satz 9. A. Es gebe zu jedem Punkt eine die Bedingung (S) erfüllende Umgebung. Dann zeigen wir zunächst:

Die Blätterzahl $B(\alpha, \mathfrak{T})$ *ist endlich.*

Wäre dies nämlich unwahr, so ließe sich zu jeder natürlichen Zahl n ein Punkt $\mathfrak{p}_n$ mit $(\alpha(\mathfrak{p}_n), \mathfrak{T}) \geqq n$ auswählen. Da alle Vielfachheiten endlich sind, ist die Menge $\{\mathfrak{p}_n\}$ unendlich und hat wegen der Kompaktheit von $\mathfrak{T}$ wenigstens einen Hp $\mathfrak{q}$. Ist dann $\mathfrak{U}$ eine der Bedingung (S) genügende Umgebung von $\mathfrak{q}$, so enthält sie einerseits unendlich viele Punkte $\mathfrak{p}_n$, während andererseits wegen der Bedingung (S) alle Punkte aus $\mathfrak{U}$ die gleiche [endliche] Vielfachheit wie $\mathfrak{q}$ haben müssen; Widerspruch!

Es sei $\mathsf{T}(\mathfrak{p})$ die Gesamtheit aller der Abbildungen aus T, bei denen $\mathfrak{p}$ und alle zu $\mathfrak{p}$ äquivalenten Punkte festbleiben. Dann ist $\mathsf{T}(\mathfrak{p})$ ein Normalteiler von T und die Faktorgruppe $\mathsf{T}/\mathsf{T}(\mathfrak{p})$ isomorph zur Gruppe aller Permutationen der zu $\mathfrak{p}$ äquivalenten Punkte, die durch T bewirkt werden. Also ist $\mathsf{T}/\mathsf{T}(\mathfrak{p})$ endlich, und zwar ist ihre Ordnung höchstens gleich $(B(\alpha, \mathfrak{T}))!$.

Ist $\mathsf{T}(\mathfrak{p}) > \mathsf{T}(\mathfrak{q})$, so ist die Ordnung von $\mathsf{T}/\mathsf{T}(\mathfrak{p})$ kleiner [mit Ausschluß der Gleichheit] als die Ordnung von $\mathsf{T}/\mathsf{T}(\mathfrak{q})$. Da die Ordnungen dieser Restklassengruppen sämtlich höchstens $(B(\alpha, \mathfrak{T}))!$, also beschränkt sind, so gibt es zu jedem Punkte $\mathfrak{p}$ wenigstens einen Punkt $\mathfrak{a}$ mit folgenden beiden Eigenschaften:

a) $\mathsf{T}(\mathfrak{a}) \leqq \mathsf{T}(\mathfrak{p})$.

b) Ist $\mathsf{T}(\mathfrak{b}) \leqq \mathsf{T}(\mathfrak{a})$, so ist $\mathsf{T}(\mathfrak{b}) = \mathsf{T}(\mathfrak{a})$.

Mithin ist jeder Punkt in wenigstens einer Menge $\mathfrak{A}_\nu$ enthalten, in der es einen Punkt $\mathfrak{a}_\nu$ gibt, so daß

$\mathfrak{a}_0$) ein Punkt $\mathfrak{p}$ dann und nur dann zu $\mathfrak{A}_\nu$ gehört, wenn $\mathsf{T}(\mathfrak{p}) \geqq \mathsf{T}(\mathfrak{a}_\nu)$ ist, und

$\mathfrak{b}_0$) aus $\mathsf{T}(\mathfrak{b}) \leqq \mathsf{T}(\mathfrak{a}_\nu)$ stets auch $\mathsf{T}(\mathfrak{b}) = \mathsf{T}(\mathfrak{a}_\nu)$ folgt.

Dann gilt:

Jedes $\mathfrak{A}_\nu$ *ist gegenüber* T *invariant und erfährt bei* T *nur endlich viele Abbildungen.*

Nach Definition von $\mathfrak{A}_\nu$ erleidet nämlich dieses bei jeder Abbildung aus $\mathsf{T}(\mathfrak{a}_\nu)$ die identische Abbildung. Es werden also die sämtlichen verschiedenen Abbildungen, die $\mathfrak{A}_\nu$ bei T erfährt, durch die endlich vielen Restklassen von $\mathsf{T}/\mathsf{T}(\mathfrak{a}_\nu)$ repräsentiert.

Ferner gilt:

Es kann nur endlich viel verschiedene $\mathfrak{A}_\nu$ *geben.*

Wäre dies nämlich unwahr, so könnte man aus jedem $\mathfrak{A}_\nu$ ein $\mathfrak{a}_\nu$ auswählen. Diese unendliche Menge muß wegen der Kompaktheit von $\mathfrak{T}$

wenigstens einen Hp c haben. Ist dann $\mathfrak{U}$ eine der Bedingung (S) genügende Umgebung von c, so gilt für jeden Punkt r aus $\mathfrak{U}$, daß $T(r) = T(c)$ ist. Also erzeugen die unendlich vielen α_ν, die in $\mathfrak{U}$ enthalten sein müssen, alle dasselbe $\mathfrak{A}_\nu$; Widerspruch!

Da jedes der $\mathfrak{A}_\nu$ bei T nur endlich viel verschiedene Abbildungen, erleidet, da weiter jeder Punkt aus $\mathfrak{T}$ in einem $\mathfrak{A}_\nu$ enthalten ist, und da es schließlich nur endlich viel verschiedene $\mathfrak{A}_\nu$ geben kann, so erfährt $\mathfrak{T}$ bei T nur endlich viel verschiedene Abbildungen, d. h. T ist endlich.

B. Es sei T endlich, $\mathfrak{p}$ irgendein Punkt aus $\mathfrak{T}$. Dann gibt es stets eine Umgebung $\mathfrak{U}$ von $\mathfrak{p}$ derart, daß α auf $\mathfrak{U}$ topologisch ist und kein von $\mathfrak{p}$ verschiedener, zu $\mathfrak{p}$ äquivalenter Punkt Hp von $\mathfrak{U}$ ist.

Sind dann $\mathfrak{U}_1, \ldots, \mathfrak{U}_m$ die sämtlichen aus $\mathfrak{U}$ durch Abbildungen aus T hervorgehenden Mengen, die $\mathfrak{p}$ nicht enthalten, so ist jedes $\mathfrak{U}_j$ offen und $\mathfrak{p}$ nicht Hp von $\mathfrak{U}_j$. Ist dann

$$\mathfrak{V} = \mathfrak{U} - [\mathfrak{U} \cap \overline{\mathfrak{U}_1 \vee \ldots \vee \mathfrak{U}_m}],$$

so ist $\mathfrak{V}$ ebenfalls eine Umgebung von $\mathfrak{p}$.

Sind weiter $\mathfrak{V}_1, \ldots, \mathfrak{V}_r$ die sämtlichen Umgebungen von $\mathfrak{p}$, in die $\mathfrak{V}$ bei Abbildungen aus T übergeht, so ist auch

$$\mathfrak{W} = \mathfrak{V}_1 \cap \ldots \cap \mathfrak{V}_r$$

eine Umgebung von $\mathfrak{p}$.

Der Beweis ist beendet, wenn gezeigt wird:

$\mathfrak{W}$ *erfüllt die Bedingung* (S).

Ist nämlich τ eine Abbildung aus T, bei der $\mathfrak{W}$ wieder in eine Umgebung von $\mathfrak{p}$ übergeht, so geht wegen $\mathfrak{W} \leqq \mathfrak{V}_j$ auch jedes $\mathfrak{V}_j$ in eine Umgebung von $\mathfrak{p}$, d. h. in ein $\mathfrak{V}_l$ über. Also ist $\tau(\mathfrak{W}) = \mathfrak{W}$, und da α auf $\mathfrak{U} \geqq \mathfrak{W}$ topologisch ist, so ist τ in $\mathfrak{W}$ die identische Abbildung.

Geht aber $\mathfrak{W}$ bei τ nicht in eine Umgebung von $\mathfrak{p}$ über, so folgt aus $\mathfrak{W} \leqq \mathfrak{V} \leqq \mathfrak{U}$, daß $\mathfrak{U}$ bei τ in eine Umgebung eines zu $\mathfrak{p}$ äquivalenten Punktes $\mathfrak{q} \neq \mathfrak{p}$ übergeht, da α in $\mathfrak{W}$, $\mathfrak{U}$, $\mathfrak{V}$ topologisch ist. Also ist $\tau(\mathfrak{U}) = \mathfrak{U}_j$ in obiger Bezeichnung und also:

$$\begin{aligned}\mathfrak{W} \cap \tau(\mathfrak{W}) \leqq \mathfrak{V} \cap \tau(\mathfrak{V}) &= \{\mathfrak{U} - [\mathfrak{U} \cap \overline{\mathfrak{U}_1 \vee \ldots \vee \mathfrak{U}_m}]\} \cap \tau(\mathfrak{V}) \\ &\leqq \{\mathfrak{U} - [\mathfrak{U} \cap \overline{\mathfrak{U}_1 \vee \ldots \vee \mathfrak{U}_m}]\} \cap \tau(\mathfrak{U}) \\ &= \{\mathfrak{U} - [\mathfrak{U} \cap \overline{\mathfrak{U}_1 \vee \ldots \vee \mathfrak{U}_m}]\} \cap \mathfrak{U}_j = 0,\end{aligned}$$

d. h. $\mathfrak{W} \cap \tau(\mathfrak{W}) = 0$.

Nicht jede diskontinuierliche Gruppe T topologischer Abbildungen eines kompakten Raumes $\mathfrak{T}$ ist endlich, wie das folgende Beispiel zeigt:

$\mathfrak{T}$ bestehe aus Punkten $\mathfrak{p}_{i,n}$ mit $i = 1, 2$ und $n = 0, 1, 2, \ldots$. Eine Teilmenge aus $\mathfrak{T}$ ist offen, wenn sie mit einem Punkte $\mathfrak{p}_{i,o}$ fast alle Punkte $\mathfrak{p}_{i,n}$

enthält. T enthalte einmal die Abbildung

$$\mathfrak{p}_{1,n} \longleftrightarrow \mathfrak{p}_{2,n},$$

weiter alle Abbildungen, bei denen für endlich viele $n > 0$ gilt

$$\mathfrak{p}_{1,n} \longleftrightarrow \mathfrak{p}_{2,n},$$

sonst aber

$$\mathfrak{p}_{i,n} \longleftrightarrow \mathfrak{p}_{i,n},$$

schließlich alle Produkte aus diesen Abbildungen.

Dann ist $\mathfrak{T}$ kompakt, T unendlich, aber, wie aus Satz 7 folgt, diskontinuierlich.

Ist die Gruppe T *topologischer Abbildungen des Raumes* $\mathfrak{T}$ *endlich,* $\mathfrak{p}$ *ein beliebiger Punkt aus* $\mathfrak{T}$ *und* n *die* [*endliche*] *Zahl der zu* $\mathfrak{p}$ *hinsichtlich* T *äquivalenten Punkte, so ist* n *ein Teiler der Ordnung von* T.

Versteht man nämlich unter $\mathsf{T}(\mathfrak{p})$ die Untergruppe aller und nur der Abbildungen aus T, die jeden zu $\mathfrak{p}$ kongruenten Punkt in Ruhe lassen, so ist $\mathsf{T}(\mathfrak{p})$ Normalteiler von T und $\mathsf{T}/\mathsf{T}(\mathfrak{p})$ isomorph zur Gruppe der Permutationen der Klasse der zu $\mathfrak{p}$ äquivalenten Punkte, die durch Abbildungen aus T bewirkt werden. Die Ordnung dieser Faktorgruppe ist aber einerseits ein Teiler der Ordnung von T, andererseits wegen der Transitivität der Permutationsgruppe ein Vielfaches von n.

(Eingegangen am 9. Februar 1931.)

Umkehrbare Ideale.[1])

Von

Werner Weber in Göttingen.

In einer Note[2]) habe ich gezeigt, daß die Theorie der Ordnungen des quadratischen Zahlkörpers unter Benutzung der Umkehrbarkeit ihrer Ideale eine sehr einfache Gestalt annimmt. Die folgende Arbeit soll diesen Gedanken zu einer Theorie der umkehrbaren Ideale in allgemeineren Ringen fortsetzen. Im Vordergrund stehen zwei Fragen: die Aufspaltung der Idealklassen beim Übergang zu weniger umfassenden Ordnungen ohne Erniedrigung des Ranges und die Beziehungen der Norm zu den konjugierten Idealen. Beide Aufgaben sind im Sonderfall des quadratischen Zahlkörpers von Dedekind[3]) gelöst worden; die zweite ist überdies für die Hauptordnung jedes endlichen algebraischen Zahlkörpers längst in demselben Sinne erledigt. Obgleich alle in diesen besonderen Fällen vorkommenden Ideale umkehrbar sind, lassen sich die alten Beweise nicht auf umkehrbare Ideale in beliebigen Körpern ausdehnen. Daß aber trotzdem ähnliche Verhältnisse auch in viel allgemeineren Ordnungen walten, wird in der nachstehenden Arbeit gezeigt.

Als angemessen erweisen sich die Dedekindsche Begriffsbildung der Ordnung eines Moduls, hier meist eines Ideals, und der hieraus sich ergebende Begriff des Bereiches derjenigen Ideale in einer Ordnung $\mathfrak{n}$, die zugleich die Ordnung $\mathfrak{n}$ „haben“[4]). Wählt man bei gegebenem $\mathfrak{n}$ unter

[1]) Von der Mathematisch-Naturwissenschaftlichen Fakultät der Universität Göttingen als Habilitationsschrift angenommen.

[2]) Bemerkungen zur arithmetischen Theorie der binären quadratischen Formen (Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse, Jahrgang 1929, S. 116—130). Auf diese Note wird fortan mit W. verwiesen.

[3]) Dirichlet-Dedekind, Vorlesungen über Zahlentheorie, 4. Aufl. (1894) (fortan mit D. 1 bezeichnet), § 187, S. 650—657 und S. 645.

[4]) Der Begriff dieser Ideale (nicht nur der Moduln von einer gewissen Ordnung) findet sich schon an einer vereinzelten Stelle in D. 1, S. 580.

diesen Idealen die umkehrbaren aus (was im quadratischen Falle und in Hauptordnungen keine besondere Einschränkung ist), so entsteht eine für die vorliegenden Aufgaben zweckmäßige Verallgemeinerung desjenigen Idealbereiches, der in W.[5]) mit $\mathfrak{B}_n$ bezeichnet war. Auf einen einzelnen solchen Bereich beziehen sich die grundlegenden Hilfssätze (§§ 1 und 2). Das wichtigste Rüstzeug bildet der Satz, daß jedes Ideal des Bereiches $\mathfrak{B}_n$ sich durch Multiplikation mit einem zu einem vorgegebenen Ideal teilerfremden Ideal in ein Hauptideal verwandeln läßt.

Die Aufspaltung der Idealklassen ergibt sich nun ohne Schwierigkeit, wenn einmal die Beziehungen zwischen den umkehrbaren und den zum „Führer" teilerfremden Idealen geklärt sind. Wie sich zunächst zeigt, besteht in der minder umfassenden Ordnung eine enge Verwandtschaft zwischen den umkehrbaren Idealen und denjenigen zum Führer teilerfremden Idealen, deren „Erweiterungsideale" bereits umkehrbar sind: es stimmen zwar nicht notwendig diese Idealgattungen selbst, wohl aber die durch sie vertretenen Idealklassen völlig überein. Damit überträgt sich die von Dedekind geschaffene Theorie der zum Führer teilerfremden Ideale sinngemäß auf die umkehrbaren. Zugleich ist gezeigt, daß in „sehr vielen" endlichen Zahlkörpern (nämlich zum mindesten in denjenigen, deren Hauptordnung nicht die Klassenzahl 1 hat) jede Ordnung vom Rang der Hauptordnung neben den Hauptidealen noch andere umkehrbare Ideale enthält. Denn die Klasse jedes Nicht-Hauptideals der Hauptordnung spaltet sich in einer beliebigen andern Ordnung vom selben Rang in eine Schar von Klassen aus Nicht-Hauptidealen auf. Nicht nur in quadratischen Zahlkörpern, sondern noch in einer großen Reihe anderer Körper gibt es somit „nichttriviale" umkehrbare Ideale.

Die vollständige Auflösung der „linearen Gleichung" $\mathfrak{m}\,\mathfrak{m}_1 = \mathfrak{M}$, in der etwa alle Buchstaben gebrochene Ideale (Moduln von endlichem Rang) bezeichnen, schließt sich bei ausreichenden Umkehrbarkeitsvoraussetzungen und -forderungen sofort an, und zwar in genauer Anlehnung an Dedekinds Behandlung des quadratischen Zahlkörpers[6]).

Unabhängig von den Ergebnissen dieses Abschnitts zeigt sich, daß die Darstellung der Norm als Produkt der konjugierten Ideale sich auf alle umkehrbaren Ideale übertragen läßt, wofern man zur Erzielung eines zweckmäßigen Normbegriffes den Grundbereich als Hauptidealring annimmt. Der nachstehende Beweis (§ 6) arbeitet wesentlich mit dem Normenmultiplikationssatz für umkehrbare Ideale einer und derselben Ordnung, der sich mittels der grundlegenden Sätze aus §§ 1 und 2 rasch beweisen läßt. Als

[5]) Vgl. Fußnote [2]).
[6]) D. 1, S. 650—655.

leichte Folgerung des Zusammenhangs zwischen Norm und Konjugierten ergibt sich aber hernach der allgemeinere Normenmultiplikationssatz für umkehrbare Ideale, deren Ordnungen denselben Quotientenkörper haben, aber nicht notwendig übereinstimmen.

Daß die Erhaltung der für die Hauptordnung gültigen Sätze trotzdem bei allgemeinen umkehrbaren Idealen ziemlich enge Grenzen hat, zeigt der letzte Teil dieser Arbeit (§ 7). Da für die Ideale eines Bereiches $\mathfrak{B}_\mathfrak{n}$ insbesondere aus Teilbarkeit Produktdarstellung folgt, so könnte man eine weitgehende Übertragung der Zerlegungseigenschaften aus den Hauptordnungen der Zahlkörper, insbesondere eine Erhaltung der eindeutigen Faktorzerlegung vermuten. Aber z. B. in endlichen algebraischen Zahlkörpern ist das Gegenteil der Fall: In keiner einzigen endlichen Ordnung $\mathfrak{n}$, die nicht die Hauptordnung ihres Quotientenkörpers ist, gestatten die Ideale des zugehörigen Bereiches $\mathfrak{B}_\mathfrak{n}$ eine eindeutige Faktorzerlegung in Ideale eines festen Bereiches. Von diesem braucht man dabei weder vorauszusetzen, daß er aus Idealen von $\mathfrak{B}_\mathfrak{n}$, noch daß er aus Primidealen besteht. Es handelt sich hier um ein Seitenstück zu dem Satz von E. Noether[7]), daß aus der eindeutigen Primfaktorzerlegung der sämtlichen Ideale in einem Ring $\mathfrak{n}$ unter anderem[8]) folgt, daß $\mathfrak{n}$ (nullteilerfrei und) im Quotientenkörper ganzabgeschlossen ist. Wie man jetzt erkennt, läßt in endlichen algebraischen Zahlkörpern schon der im allgemeinen engere Bereich $\mathfrak{B}_\mathfrak{n}$ eine Entscheidung in diesem Sinne zu. Fällt die Entscheidung in einem Einzelfall bejahend aus, so ist damit nachträglich gezeigt, daß $\mathfrak{B}_\mathfrak{n}$ in diesem Fall eben doch keine echte Untermenge des Bereiches aller Ideale in $\mathfrak{n}$ ist.

§ 1.

Umkehrbare Ideale in Körperordnungen.

In dieser Arbeit bedeutet $\mathfrak{R}$ einen nullteilerfreien Ring mit Einheitselement $1 \neq 0$, K einen $\mathfrak{R}$ umfassenden kommutativen Körper. „*Modul*“ bedeutet $\mathfrak{R}$-Modul[9]) aus K. Der *Quotient* $\frac{\mathfrak{a}}{\mathfrak{b}}$ zweier Moduln $\mathfrak{a}, \mathfrak{b}$ wird als die Gesamtheit aller Elemente α von K mit $\alpha\,\mathfrak{b} \leqq \mathfrak{a}$[10]) erklärt[11]) und ist selbst

---

[7]) Abstrakter Aufbau der Idealtheorie in algebraischen Zahl- und Funktionenkörpern, Math. Annalen 96 (1927), S. 26—61 (künftig mit N. bezeichnet), § 9.

[8]) Daß der Satz von E. Noether mit dieser Teilaussage noch lange nicht erschöpft ist, braucht wohl nicht besonders hervorgehoben zu werden.

[9]) Vgl. etwa N., S. 30.

[10]) Hier seien ein für allemal die im folgenden benutzten mengentheoretischen Bezeichnungen zusammengestellt: $=$ bedeutet Übereinstimmung, $\neq$ Verschiedenheit, $\leqq$ Untermenge, $<$ echte Untermenge, $\geqq$ Obermenge, $\in$ das Enthaltensein, $\cap$ den Durchschnitt.

[11]) Vgl. D. 1, S. 504.

ein Modul. Ist $\mathfrak{m}$ ein Modul, so wird der Modul $\frac{\mathfrak{m}}{\mathfrak{m}}$ mit $\mathfrak{m}^0$ bezeichnet; $\mathfrak{m}^0$, die „*Ordnung*“[12]) von $\mathfrak{m}$, ist offenbar ein $\mathfrak{R}$ umfassender Unterring von K. Jeder $\mathfrak{R}$ umfassende Unterring von K ist in diesem Sinne eine Ordnung, nämlich seine eigene. Der Begriff der Ordnung deckt sich also mit dem üblichen Begriff[13]) der $\mathfrak{R}$-Ordnung aus K und wird im folgenden, außer in Satz 25, stets in diesem Sinne gebraucht. Ist $\mathfrak{m}$ ein Modul, so bezeichnet $\mathfrak{m}^{-1}$ den Modul $\frac{\mathfrak{m}^0}{\mathfrak{m}}$.[14])

Die Begriffe $\mathfrak{m}^0$ und $\mathfrak{m}^{-1}$ hängen nach dieser Erklärung zunächst noch von der Wahl des Körpers K ab. Liegen aber im Durchschnitt einer Menge von Körpern K der vom Nullmodul verschiedene Modul $\mathfrak{m}$ und alle Quotienten $\frac{\alpha}{\beta}$, wo α und β dem Modul $\mathfrak{m}$ angehören und $\beta \neq 0$ ist, so besteht jene Abhängigkeit nur scheinbar; denn offensichtlich ist jedes Element von $\mathfrak{m}^0$ und infolgedessen auch jedes Element von $\mathfrak{m}^{-1}$ einer jener Quotienten $\frac{\alpha}{\beta}$. In diesem Fall braucht also nicht angegeben zu werden, auf welchen Körper K sich der Begriff $\mathfrak{m}^0$ oder $\mathfrak{m}^{-1}$ bezieht.

Definition 1 (Dedekind[14])). *Der Modul* $\mathfrak{m}$ *heißt umkehrbar, wenn*

$$\mathfrak{m}^{-1}\mathfrak{m} = \mathfrak{m}^0$$

ist.

Jeder umkehrbare Modul ist also vom Nullmodul verschieden. — Die Entscheidung über Umkehrbarkeit ist offenbar im vorstehend erklärten Sinne von der Wahl des Körpers K unabhängig.

Nach Dedekind[15]) gelten die folgenden drei Sätze:

Satz 1 (Dedekind). *Ein Modul* $\mathfrak{m}$ *von der Ordnung* $\mathfrak{n}$ *ist dann und nur dann umkehrbar, wenn es einen Modul* $\mathfrak{a}$ *mit* $\mathfrak{m}\mathfrak{a} = \mathfrak{n}$ *gibt.*

Satz 2 (Dedekind). *Ist der Modul* $\mathfrak{m}$ *umkehrbar, so ist* $\mathfrak{m}^{-1}$ *ebenfalls umkehrbar und*

$$(\mathfrak{m}^{-1})^0 = \mathfrak{m}^0.$$

Satz 3 (Dedekind). *Sind* $\mathfrak{a}$ *und* $\mathfrak{b}$ *umkehrbare Moduln, so ist auch* $\mathfrak{a}\mathfrak{b}$ *umkehrbar und*

$$\mathfrak{a}^0\mathfrak{b}^0 = (\mathfrak{a}\mathfrak{b})^0.$$

Definition 2. *Ist* $\mathfrak{n}$ *eine Ordnung, so bedeutet* $\mathfrak{B}_\mathfrak{n}$ *den Bereich aller derjenigen Ideale* $\mathfrak{a}$ *im Ring* $\mathfrak{n}$, *die*

1. (als Moduln betrachtet) *umkehrbar sind,*
2. *die Eigenschaft* $\mathfrak{a}^0 = \mathfrak{n}$ *haben.*

---

[12]) Vgl. D. 1, S. 505.
[13]) Vgl. N., S. 31.
[14]) Vgl. D. 1, S. 506.
[15]) Vgl. D. 1, S. 506 f.

Der Bereich $\mathfrak{B}_\mathfrak{n}$ ist nach dem Obigen bei gegebenem $\mathfrak{n}$ derselbe in bezug auf alle $\mathfrak{n}$ umfassenden Körper K.

Satz 4. *Ist* $\mathfrak{n}$ *eine Ordnung,* $\mathfrak{a}$ *Element von* $\mathfrak{B}_\mathfrak{n}$, $\mathfrak{b}$ *Ideal in* $\mathfrak{n}$, $\mathfrak{m}$ *ein Modul und* $\mathfrak{a} = \mathfrak{b}\mathfrak{m}$, *so gehört auch* $\mathfrak{b}$ *dem Bereich* $\mathfrak{B}_\mathfrak{n}$ *an.*

Beweis. Es ist $\mathfrak{a}\mathfrak{a}^{-1} = \mathfrak{n}$, also $\mathfrak{b}(\mathfrak{m}\mathfrak{a}^{-1}) = \mathfrak{n}$. Für jedes Element α von K mit $\alpha\mathfrak{b} \leqq \mathfrak{b}$ ist $\alpha \in \alpha\mathfrak{n} = (\alpha\mathfrak{b})(\mathfrak{m}\mathfrak{a}^{-1}) \leqq \mathfrak{b}(\mathfrak{m}\mathfrak{a}^{-1}) = \mathfrak{n}$; es ist also $\mathfrak{b}^0 \leqq \mathfrak{n}$, also (wegen der Idealeigenschaft von $\mathfrak{b}$) $\mathfrak{b}^0 = \mathfrak{n}$. Wegen $\mathfrak{b}(\mathfrak{m}\mathfrak{a}^{-1}) = \mathfrak{n}$ ist $\mathfrak{b}$ nach Satz 1 umkehrbar.

Satz 5. 1. *Sind* $\mathfrak{a}$ *und* $\mathfrak{b}$ *Ideale in der Ordnung* $\mathfrak{n}$, *das Ideal* $\mathfrak{b}$ *überdies zu* $\mathfrak{B}_\mathfrak{n}$ *gehörig und ist* $\mathfrak{a}$ *durch* $\mathfrak{b}$ *teilbar, so gibt es in* $\mathfrak{n}$ *ein Ideal* $\mathfrak{c}$ *mit*

$$\mathfrak{a} = \mathfrak{b}\mathfrak{c}. \tag{1}$$

2. *Dieses* $\mathfrak{c}$ *ist eindeutig bestimmt; es ist nämlich* $\mathfrak{c} = \mathfrak{b}^{-1}\mathfrak{a}$. *Liegt* $\mathfrak{a}$ *insbesondere in* $\mathfrak{B}_\mathfrak{n}$, *so liegt auch* $\mathfrak{c}$ *in* $\mathfrak{B}_\mathfrak{n}$.

Beweis. Der Nachweis von 1. geschieht wörtlich ebenso wie der des Hilfssatzes 1 von W. (S. 120), wo der Sonderfall des quadratischen Zahlkörpers vorliegt. Auf 2. schließt man folgendermaßen: Aus (1) folgt

$$\mathfrak{c} = \mathfrak{n}\mathfrak{c} = \mathfrak{b}^{-1}\mathfrak{b}\mathfrak{c} = \mathfrak{b}^{-1}\mathfrak{a};$$

liegt $\mathfrak{a}$ in $\mathfrak{B}_\mathfrak{n}$, so ist also $\mathfrak{c}$ nach Satz 2 und Satz 3 umkehrbar und

$$\mathfrak{c}^0 = (\mathfrak{b}^{-1})^0\mathfrak{a}^0 = \mathfrak{b}^0\mathfrak{n} = \mathfrak{n}\mathfrak{n} = \mathfrak{n}.$$

Satz 3 und Satz 5 zeigen schon, daß die Ideale des Bereiches $\mathfrak{B}_\mathfrak{n}$ gewisse wesentliche Eigenschaften, insbesondere nämlich die Gruppeneigenschaft ihrer „Klassen", mit den Idealen in der Hauptordnung eines endlichen algebraischen Zahlkörpers gemein haben. In dem besonderen Falle nun, daß $\mathfrak{R}$ der Ring der ganzen rationalen Zahlen und K ein endlicher algebraischer Zahlkörper ist, ist neben der für alle Ringe gebräuchlichen Einteilung in Idealklassen noch eine schärfere üblich, bei der zur Hauptklasse nur diejenigen Hauptideale gehören, die durch Elemente von positiver Norm erzeugt werden können. Im allgemeinen Fall gibt es diesen Begriff nicht. Wenn der angegebene Sonderfall vorliegt und man sich überdies für die schärfere Klasseneinteilung entscheidet, so werden im folgenden einige Zusätze nötig; sie sind in fetten eckigen Klammern beigefügt.

Satz 6. *Ist* $\mathfrak{n}$ *eine Ordnung,* $\mathfrak{a}$ *Element von* $\mathfrak{B}_\mathfrak{n}$, *so kann* $\mathfrak{a}$ *durch Multiplikation mit einem Ideal aus* $\mathfrak{B}_\mathfrak{n}$ *in ein vom Nullideal verschiedenes Hauptideal* $\alpha\mathfrak{n}$ *verwandelt werden, wo* α *in* $\mathfrak{n}$ *liegt* [*und positive Norm hat*].

Beweis. Ist α irgendein von Null verschiedenes Element von $\mathfrak{a}$ [mit positiver Norm], so ist $\alpha\mathfrak{n} \subseteq \mathfrak{a}$. Da K Körper ist, so ist das Ideal $\alpha\mathfrak{n}$ offensichtlich umkehrbar, also Element von $\mathfrak{B}_\mathfrak{n}$. Nach Satz 5 liegt also in $\mathfrak{B}_\mathfrak{n}$ ein Ideal $\mathfrak{x}$ mit

$$\alpha\mathfrak{n} = \mathfrak{a}\mathfrak{x}.$$

Ist $\mathfrak{n}$ eine Ordnung, so bedeutet künftig eine *Idealklasse* von $\mathfrak{n}$, wie üblich, die Gesamtheit aller derjenigen Ideale $\mathfrak{b}$ in $\mathfrak{n}$, die zu einem festen Ideal $\mathfrak{a}$ in $\mathfrak{n}$ in einer Beziehung von der Gestalt

$$\eta_1\mathfrak{a} = \eta_2\mathfrak{b}$$

stehen, wo η_1 und η_2 nichtverschwindende Elemente von $\mathfrak{n}$ sind $\left[\text{und } \frac{\eta_1}{\eta_2} \text{ positive Norm hat}\right]$. Im Sinne dieser Klasseneinteilung wird das Wort *„äquivalent“* gebraucht.

Jedes zu einem Ideal aus $\mathfrak{B}_\mathfrak{n}$ äquivalente Ideal in $\mathfrak{n}$ liegt offenbar ebenfalls in $\mathfrak{B}_\mathfrak{n}$. In diesem Sinne kann man von *„Idealklassen aus* $\mathfrak{B}_\mathfrak{n}$“ reden.

Sind A und B Idealklassen von $\mathfrak{n}$, so sind alle Produkte aus je einem Element von A und einem Element von B äquivalent; die durch sie vertretene Klasse heißt, wie üblich, das *Produkt* der Klassen A und B. Aus Satz 3 und Satz 6 ergibt sich, daß die Idealklassen aus $\mathfrak{B}_\mathfrak{n}$ gegenüber dieser Multiplikation eine Abelsche Gruppe bilden, deren Einheitselement die Klasse der vom Nullideal verschiedenen [und aus Elementen von positiver Norm abgeleiteten] Hauptideale ist.

Ohne weitere Einschränkungen des Ordnungsbegriffes bleibt auch der folgende, für Hauptordnungen endlicher algebraischer Zahlkörper von Dedekind[16]) bewiesene, für diese Arbeit entbehrliche Satz bestehen:

Satz 7. *Sind* $\mathfrak{c}_1, \ldots, \mathfrak{c}_r$ $(r \geqq 1)$ *Ideale in der Ordnung* $\mathfrak{n}$, *das Ideal* $\mathfrak{a}$ *Element von* $\mathfrak{B}_\mathfrak{n}$ *und ist* $\mathfrak{a}$ *durch keines der Ideale* $\mathfrak{c}_i$ $(i = 1, \ldots, r)$ *teilbar, so gibt es in* $\mathfrak{a}$ *ein Element* [*von positiver Norm*], *das in keinem der Ideale* $\mathfrak{c}_i$ *liegt.*

Beweis wie bei W., Hilfssatz 2 (S. 120 f.), wobei der Schlußteil, der sich auf die positive Norm bezieht, im allgemeinen Fall zu streichen ist. Das dort angewandte Beweisverfahren ist gerade das Dedekindsche[16]).

§ 2.

Umkehrbare Ideale in Ringen mit eingeschränktem Doppelkettensatz.

In diesem Paragraphen bedeutet $\mathfrak{n}$ stets eine Ordnung, in der für die Ideale der Teilerkettensatz[17]) und der Vielfachenkettensatz mod den vom

[16]) D. 1, S. 558.

[17]) Zu diesem Begriff s. N., S. 26.

Nullideal verschiedenen Idealen[17]) gelten[18]). Die Idealtheorie des Ringes $\mathfrak{n}$ ergibt sich aus N.[19]), § 7; insbesondere läßt sich in $\mathfrak{n}$ jedes vom Nullideal verschiedene Ideal als Produkt[20]) paarweise teilerfremder, vom Einheitsideal verschiedener Primärideale darstellen, und zwar bis auf die Reihenfolge der Faktoren eindeutig.

Satz 8. *Ist* $\mathfrak{a}$ *ein Element von* $\mathfrak{B}_\mathfrak{n}$ *und ist* $\mathfrak{b}$ *ein beliebiges vom Nullideal verschiedenes Ideal in* $\mathfrak{n}$, *so gibt es in* $\mathfrak{B}_\mathfrak{n}$ *ein zu* $\mathfrak{b}$ *teilerfremdes Ideal* $\mathfrak{x}$ *von der Beschaffenheit, daß* $\mathfrak{a}\mathfrak{x}$ *ein Hauptideal* $\alpha\mathfrak{n}$ *ist, wo* α *in* $\mathfrak{n}$ *liegt* [*und positive Norm hat*].

Beweis. Ist $\mathfrak{a}$ das Einheitsideal, so leistet $\mathfrak{x} = \mathfrak{n}$ das Gewünschte. Es sei nunmehr $\mathfrak{a}$ vom Einheitsideal verschieden, also von der Gestalt

$$\mathfrak{a} = \mathfrak{q}_1 \cdots \mathfrak{q}_r \qquad (r \geqq 1),$$

wo die $\mathfrak{q}_i$ $(i = 1, \ldots, r)$ paarweise teilerfremde, vom Einheitsideal verschiedene Primärideale sind; $\mathfrak{p}_1, \ldots, \mathfrak{p}_r$ seien die zugehörigen Primideale. Nach Satz 4 gehört jedes $\mathfrak{q}_i$ dem Bereich $\mathfrak{B}_\mathfrak{n}$ an. In der Darstellung von $\mathfrak{b}$ als Produkt[20]) paarweise teilerfremder, vom Einheitsideal verschiedener Primärideale sei $\mathfrak{q}$ das Produkt derjenigen, deren zugehörige Primideale unter den $\mathfrak{p}_i$ vorkommen, $\mathfrak{r}$ das Produkt der übrigen.

Für $i = 1, \ldots, r$ ist $\mathfrak{q}_i \mathfrak{p}_i \subset \mathfrak{q}_i$; denn aus $\mathfrak{q}_i \mathfrak{p}_i = \mathfrak{q}_i = \mathfrak{q}_i \mathfrak{n}$ würde nach Satz 5, 2. folgen $\mathfrak{p}_i = \mathfrak{n}$. Es sei α_i ein Element von $\mathfrak{q}_i$, das nicht in $\mathfrak{q}_i \mathfrak{p}_i$ liegt. Da die $\mathfrak{q}_i \mathfrak{p}_i$ paarweise teilerfremd sind, so gibt es in $\mathfrak{n}$ für $i = 1, \ldots, r$ ein η_i mit

$$\eta_i \equiv \alpha_i \pmod{\mathfrak{q}_i \mathfrak{p}_i},$$
$$\eta_i \equiv 0 \pmod{\mathfrak{q}_k \mathfrak{p}_k} \quad \text{für} \quad k \neq i.$$

Jedes η_i liegt in $\mathfrak{q}_i$, aber nicht in $\mathfrak{q}_i \mathfrak{p}_i$. Da die Ideale

$$\mathfrak{b}_i = (\mathfrak{q}_1 \mathfrak{p}_1) \cdots (\mathfrak{q}_{i-1} \mathfrak{p}_{i-1})(\mathfrak{q}_{i+1} \mathfrak{p}_{i+1}) \cdots (\mathfrak{q}_r \mathfrak{p}_r)$$

in ihrer Gesamtheit teilerfremd sind, so ist etwa

$$\delta_1 + \cdots + \delta_r = 1, \qquad \delta_i \in \mathfrak{b}_i.$$

Jedes δ_i ist durch alle $\mathfrak{p}_k$ mit $k \neq i$ teilbar; also ist kein δ_i durch $\mathfrak{p}_i$ teilbar, da sonst 1 durch $\mathfrak{p}_i$ teilbar wäre. Da endlich $(\mathfrak{q}_1 \mathfrak{p}_1) \cdots (\mathfrak{q}_r \mathfrak{p}_r)$ zu

---

[18]) Falls diese Bedingungen in $\mathfrak{R}$ erfüllt sind, so sind sie in jeder endlichen Ordnung erfüllt. Vgl. Grell, Zur Theorie der Ordnungen in algebraischen Zahl- und Funktionenkörpern, Math. Annalen 97 (1927), S. 524—558 (fortan mit G. angeführt), § 1, Satz 1. Insbesondere fallen also unter § 2 die endlichen Ordnungen über einem Hauptidealring, z. B. in Zahlkörpern und Funktionenkörpern einer Veränderlichen.

[19]) Vgl. Fußnote [7]).

[20]) Ein leeres Produkt von Idealen soll hier das Einheitsideal bedeuten.

$\mathfrak{r}$ teilerfremd ist, so gibt es in $\mathfrak{n}$ ein α mit

$$\alpha \equiv \sum_{i=1}^{r} \delta_i \eta_i \pmod{(\mathfrak{q}_1 \mathfrak{p}_1) \cdots (\mathfrak{q}_r \mathfrak{p}_r)},$$
$$\alpha \equiv 1 \pmod{\mathfrak{r}}.$$

Da jedes δ_k mit $k \neq i$ in $\mathfrak{q}_i \mathfrak{p}_i$ liegt, $\delta_i \eta_i$ aber nicht, so liegt auch $\sum_{i=1}^{r} \delta_i \eta_i$ und damit α in keinem $\mathfrak{q}_i \mathfrak{p}_i$. Insbesondere ist also $\alpha \neq 0$. [α kann aber sogar mit positiver Norm gewählt werden. Denn ist l eine natürliche Zahl, die sowohl in $(\mathfrak{q}_1 \mathfrak{p}_1) \cdots (\mathfrak{q}_r \mathfrak{p}_r)$ wie auch in $\mathfrak{r}$ vorkommt, so genügt $\alpha + ml$ für jede natürliche Zahl m ebenfalls den beiden vorstehenden Kongruenzen, hat aber für große m positive Norm.]

Da jedes δ_i in $\mathfrak{q}_1 \cdots \mathfrak{q}_{i-1} \mathfrak{q}_{i+1} \cdots \mathfrak{q}_r$, jedes η_i in $\mathfrak{q}_i$ liegt, so liegt $\sum_{i=1}^{r} \delta_i \eta_i$ und daher auch α in $\mathfrak{q}_1 \cdots \mathfrak{q}_r = \mathfrak{a}$. Nach Satz 5 enthält $\mathfrak{B}_\mathfrak{n}$ ein Ideal $\mathfrak{x}$ mit

$$\alpha \mathfrak{n} = \mathfrak{a} \mathfrak{x}.$$

$\mathfrak{x}$ ist zu dem Bestandteil $\mathfrak{q}$ von $\mathfrak{b}$ teilerfremd; denn sonst wäre $\mathfrak{x}$ durch ein $\mathfrak{p}_i$, also $\alpha \mathfrak{n}$ durch ein $\mathfrak{q}_i \mathfrak{p}_i$ teilbar. Da $\alpha \mathfrak{n}$ und damit $\mathfrak{x}$ auch zu $\mathfrak{r}$ teilerfremd ist, so ist $\mathfrak{x}$ zu $\mathfrak{b}$ teilerfremd[21]).

Satz 9. *Ist $\mathfrak{a}$ Element von $\mathfrak{B}_\mathfrak{n}$ und ist $\mathfrak{b}$ ein beliebiges vom Nullideal verschiedenes Ideal in $\mathfrak{n}$, so gibt es in $\mathfrak{n}$ ein von Null verschiedenes Element α [von positiver Norm], für welches*

$$\mathfrak{a} \mathfrak{b} + \alpha \mathfrak{n} = \mathfrak{a}$$

ist.

Beweis. In den Bezeichnungen des Satzes 8 ist

$$\mathfrak{a} \mathfrak{b} + \alpha \mathfrak{n} = \mathfrak{a} \mathfrak{b} + \mathfrak{a} \mathfrak{x} = \mathfrak{a} (\mathfrak{b} + \mathfrak{x}) = \mathfrak{a} \mathfrak{n} = \mathfrak{a}.$$

---

[21]) Die Grundgestalt dieses Beweises findet sich bei Hecke (Vorlesungen über die Theorie der algebraischen Zahlen (1923), § 26, Satz 74), wo freilich eindeutige Primfaktorzerlegung der Ideale vorliegt und daher alles viel einfacher wird. In der obigen Numerierung schließt Hecke zunächst auf Satz 9 und von diesem bei späterem Bedarf auf Satz 8. Der Grundgedanke des Beweises ist aber auch schon in dem oben angedeuteten Dedekindschen Beweis des Satzes 7 enthalten. In der Tat kann man unter Benutzung des Satzes 7 den Beweis des Satzes 8 mittels der folgenden Schlußweise stark abkürzen: Da $\mathfrak{a}$ durch keines der Ideale $\mathfrak{q}_1 \mathfrak{p}_1, \ldots, \mathfrak{q}_r \mathfrak{p}_r$ teilbar ist, so gibt es in $\mathfrak{a}$ ein Element β, das in keinem dieser Ideale liegt. Die beiden gleichzeitigen Kongruenzen

$$\alpha \equiv \beta \pmod{(\mathfrak{q}_1 \mathfrak{p}_1) \cdots (\mathfrak{q}_r \mathfrak{p}_r)},$$
$$\alpha \equiv 1 \pmod{\mathfrak{r}}$$

haben eine gemeinsame Lösung α [von positiver Norm]; diese leistet das Gewünschte. — Auch Dedekind (D. 1, S. 558 f.) verwendet Satz 7 zum Beweise der Sätze 9 und 8, benutzt aber dabei an einer Stelle den dort, aber nicht hier allgemein gültigen Satz, daß jedes vom Nullideal verschiedene Ideal nur endlich viele Teiler hat.

Satz 10. *Ist* $\mathfrak{c}$ *ein vom Nullideal verschiedenes Ideal in* $\mathfrak{n}$, *so gibt es in jeder Idealklasse aus* $\mathfrak{B}_\mathfrak{n}$ *ein zu* $\mathfrak{c}$ *teilerfremdes Ideal.*

Beweis. Satz 8 besagt gerade, daß es in der Inversen jeder Idealklasse aus $\mathfrak{B}_\mathfrak{n}$ ein zu $\mathfrak{c}$ teilerfremdes Ideal gibt.

§ 3.

Die Aufspaltung.

In diesem und dem folgenden Paragraphen bedeuten $\mathfrak{n}$ und $\mathfrak{N}$ Ordnungen, in denen der Teilerkettensatz und der Vielfachenkettensatz mod den vom Nullideal verschiedenen Idealen gelten. Es sei $\mathfrak{n} \subseteq \mathfrak{N}$. Es wird ferner vorausgesetzt, daß der Führer[22]) $\mathfrak{k}$ von $\mathfrak{n}$ nach $\mathfrak{N}$ vom Nullideal verschieden ist.

Von Dedekind und Grell[23]) werden die folgenden beiden (in viel allgemeineren Ringen gültigen) Sätze übernommen:

Satz 11 (Dedekind-Grell). *Ist* $\mathfrak{A}$ *ein zu* $\mathfrak{k}$ *teilerfremdes Ideal in* $\mathfrak{N}$, *so ist* $\mathfrak{a} = \mathfrak{A} \cap \mathfrak{n}$ *ein zu* $\mathfrak{k}$ *teilerfremdes Ideal in* $\mathfrak{n}$ (das „*Verengungsideal* von $\mathfrak{A}$ in $\mathfrak{n}$"), *und es ist* $\mathfrak{A} = \mathfrak{a}\mathfrak{N}$.

Satz 12 (Dedekind-Grell). *Ist* $\mathfrak{a}$ *ein zu* $\mathfrak{k}$ *teilerfremdes Ideal in* $\mathfrak{n}$, *so ist* $\mathfrak{A} = \mathfrak{a}\mathfrak{N}$ *ein zu* $\mathfrak{k}$ *teilerfremdes Ideal in* $\mathfrak{N}$ (das „*Erweiterungsideal* von $\mathfrak{a}$ in $\mathfrak{N}$"), *und es ist* $\mathfrak{a} = \mathfrak{A} \cap \mathfrak{n}$.

Es gilt nun (ebenfalls in viel allgemeineren Ringen) der

Satz 13. *Es sei* $\mathfrak{a}$ *ein zu* $\mathfrak{k}$ *teilerfremdes Ideal in* $\mathfrak{n}$; *das Erweiterungsideal von* $\mathfrak{a}$ *in* $\mathfrak{N}$ *habe die Ordnung* $\mathfrak{N}$. *Dann hat* $\mathfrak{a}$ *die Ordnung* $\mathfrak{n}$.

Beweis[24]). Die Ordnung von $\mathfrak{a}$ werde mit $\mathfrak{n}_1$ bezeichnet; dann ist $\mathfrak{n}_1 \supseteq \mathfrak{n}$. Für jedes Element α von K mit $\alpha\mathfrak{a} \subseteq \mathfrak{a}$ ist $\alpha(\mathfrak{a}\mathfrak{N}) \subseteq \mathfrak{a}\mathfrak{N}$, $\alpha \in (\mathfrak{a}\mathfrak{N})^0 = \mathfrak{N}$; demnach ist $\mathfrak{n}_1 \subseteq \mathfrak{N}$. Wegen

$$\mathfrak{n} = \mathfrak{a} + \mathfrak{k}$$

---

[22]) Sind $\mathfrak{o}$ und $\mathfrak{O}$ Ordnungen und ist $\mathfrak{o} \subseteq \mathfrak{O}$, so heißt der Modulquotient $\frac{\mathfrak{o}}{\mathfrak{O}}$, wie üblich, der Führer von $\mathfrak{o}$ nach $\mathfrak{O}$. Er ist bekanntlich ein Ideal in $\mathfrak{o}$ und in $\mathfrak{O}$.

[23]) In der (im folgenden mit D. 2 bezeichneten) Arbeit „Über die Anzahl der Ideal-Classen in den verschiedenen Ordnungen eines endlichen Körpers" (Festschrift zur Saecularfeier des Geburtstages von Carl Friedrich Gauss dargebracht vom Herzoglichen Collegium Carolinum zu Braunschweig, 1877; auch: Gesammelte mathematische Werke 1 (1930), S. 105–157) (§ 5) beweist Dedekind die beiden Sätze für den Fall zweier Ordnungen eines endlichen algebraischen Zahlkörpers, von denen eine die Hauptordnung ist und die andere denselben Rang hat. Die Beweise gelten aber im vorliegenden allgemeineren Fall unverändert. Die Sätze wurden dann von Grell in der Arbeit „Beziehungen zwischen den Idealen verschiedener Ringe" (Math. Annalen 97 (1927), S. 490–523) (S. 516 f.) in verallgemeinerter Fassung bewiesen.

[24]) Vgl. den Beweis des entsprechenden Satzes in D. 2, § 4. Der dortige Beweis für Zahlkörper konnte hier in der Hauptsache beibehalten werden.

ist

$$\mathfrak{n}_1 = \mathfrak{n}\mathfrak{n}_1 = \mathfrak{a}\mathfrak{n}_1 + \mathfrak{k}\mathfrak{n}_1 \subseteq \mathfrak{a} + \mathfrak{k}\mathfrak{N} = \mathfrak{a} + \mathfrak{k} = \mathfrak{n},$$

wegen $\mathfrak{n}_1 \supseteq \mathfrak{n}$ also

$$\mathfrak{n}_1 = \mathfrak{n}.$$

Satz 14. *Ist $\mathfrak{A}$ ein zu $\mathfrak{k}$ teilerfremdes Ideal aus dem Bereich $\mathfrak{B}_\mathfrak{N}$, so gehört das Verengungsideal von $\mathfrak{A}$ in $\mathfrak{n}$ zum Bereich $\mathfrak{B}_\mathfrak{n}$.*

Beweis. Nach Satz 11 ist das Ideal $\mathfrak{a} = \mathfrak{A} \cap \mathfrak{n}$ zu $\mathfrak{k}$ teilerfremd und

$$\mathfrak{A} = \mathfrak{a}\mathfrak{N}.$$

Nach Satz 13 hat also $\mathfrak{a}$ die Ordnung $\mathfrak{n}$. Wegen der Teilerfremdheit von $\mathfrak{a}$ und $\mathfrak{k}$ läßt sich das System der beiden gleichzeitigen Kongruenzen

$$\alpha \equiv 0 \pmod{\mathfrak{a}},$$
$$\alpha \equiv 1 \pmod{\mathfrak{k}}$$

durch ein Element α des Ringes $\mathfrak{n}$ lösen. Ist $\alpha = 0$, so ist $1 \in \mathfrak{k}$, $\mathfrak{n} = \mathfrak{k} = \mathfrak{N}$ und die Behauptung klar. Es sei also $\alpha \neq 0$. Nach Satz 5 liegt in $\mathfrak{B}_\mathfrak{N}$ ein Ideal $\mathfrak{C}$ mit

$$\alpha\mathfrak{N} = \mathfrak{A}\mathfrak{C}.$$

Wegen $\alpha\mathfrak{N} + \mathfrak{k} = \mathfrak{N}$ ist auch $\mathfrak{C}$ zu $\mathfrak{k}$ teilerfremd. Nach Satz 11 ist also das Verengungsideal $\mathfrak{c}$ von $\mathfrak{C}$ in $\mathfrak{n}$ zu $\mathfrak{k}$ teilerfremd und genügt der Gleichung

$$\mathfrak{C} = \mathfrak{c}\mathfrak{N}.$$

Das Ideal $\mathfrak{a}\mathfrak{c}$ ist zu $\mathfrak{k}$ teilerfremd; es ist

$$(\mathfrak{a}\mathfrak{c})\mathfrak{N} = \mathfrak{A}\mathfrak{C} = \alpha\mathfrak{N}.$$

Andrerseits ist auch $\alpha\mathfrak{n}$ zu $\mathfrak{k}$ teilerfremd und

$$(\alpha\mathfrak{n})\mathfrak{N} = \alpha\mathfrak{N}.$$

Nach Satz 12 ist also

$$\mathfrak{a}\mathfrak{c} = \alpha\mathfrak{N} \cap \mathfrak{n} = \alpha\mathfrak{n},$$
$$\mathfrak{a}\left(\frac{1}{\alpha}\mathfrak{c}\right) = \mathfrak{n} = \mathfrak{a}^0.$$

Somit ist $\mathfrak{a}$ nach Satz 1 auch umkehrbar, liegt also in $\mathfrak{B}_\mathfrak{n}$.

Durchläuft $\mathfrak{a}$ die Ideale einer Klasse a von $\mathfrak{n}$, so durchläuft $\mathfrak{a}\mathfrak{N}$ offenbar eine Schar zueinander äquivalenter Ideale in $\mathfrak{N}$. Die durch diese „Erweiterungsideale“ vertretene Klasse wird sinngemäß als die *Erweiterungsidealklasse* von a bezeichnet. Gehören die Ideale von a zu $\mathfrak{B}_\mathfrak{n}$, so gehören die Ideale der Erweiterungsidealklasse von a nach Satz 3 zu $\mathfrak{B}_\mathfrak{N}$.

Satz 15. *Die Idealklassen aus $\mathfrak{B}_\mathfrak{n}$ sind genau diejenigen Idealklassen von $\mathfrak{n}$, die durch die Verengungsideale der zu $\mathfrak{k}$ teilerfremden Ideale aus $\mathfrak{B}_\mathfrak{N}$ vertreten werden.*

Beweis. 1. Es sei a eine Klasse aus $\mathfrak{B}_\mathfrak{n}$. Nach Satz 10 enthält a ein zu $\mathfrak{f}$ teilerfremdes Ideal $\mathfrak{a}$. Alsdann liegt das Ideal $\mathfrak{A} = \mathfrak{a}\mathfrak{N}$ in $\mathfrak{B}_\mathfrak{N}$, und nach Satz 12 ist $\mathfrak{A}$ zu $\mathfrak{f}$ teilerfremd und

$$\mathfrak{a} = \mathfrak{A} \cap \mathfrak{n}.$$

2. Ist umgekehrt $\mathfrak{A}$ ein zu $\mathfrak{f}$ teilerfremdes Ideal aus $\mathfrak{B}_\mathfrak{N}$, so liegt nach Satz 14 das Verengungsideal $\mathfrak{a}$ von $\mathfrak{A}$, also auch jedes zu $\mathfrak{a}$ äquivalente Ideal in $\mathfrak{B}_\mathfrak{n}$.

Definition 3. *Die Gesamtheit derjenigen zu $\mathfrak{f}$ teilerfremden Ideale in $\mathfrak{n}$, deren Erweiterungsideale in $\mathfrak{N}$ vom Nullideal verschiedene Hauptideale sind [und durch Elemente von positiver Norm erzeugt werden können], wird mit $\mathfrak{G}(\mathfrak{n}; \mathfrak{N})$ bezeichnet.*

Satz 16. *Jedes Ideal aus $\mathfrak{G}(\mathfrak{n}; \mathfrak{N})$ gehört zu $\mathfrak{B}_\mathfrak{n}$.*

Beweis. Gehört $\mathfrak{a}$ zu $\mathfrak{G}(\mathfrak{n}; \mathfrak{N})$, so ist $\mathfrak{a}$ zu $\mathfrak{f}$ teilerfremd, also nach Satz 12

$$\mathfrak{a} = \mathfrak{a}\mathfrak{N} \cap \mathfrak{n}$$

und $\mathfrak{a}\mathfrak{N}$ zu $\mathfrak{f}$ teilerfremd. Nach Definition 3 ist aber $\mathfrak{a}\mathfrak{N}$ ein vom Nullideal verschiedenes Hauptideal, also Element von $\mathfrak{B}_\mathfrak{N}$. Satz 14 lehrt also die Behauptung.

Definition 4. *Die Mächtigkeit der Menge der in $\mathfrak{G}(\mathfrak{n}; \mathfrak{N})$ vertretenen Idealklassen von $\mathfrak{n}$ wird mit $g(\mathfrak{n}; \mathfrak{N})$ bezeichnet.*

Satz 17. 1. *Im Sinne des Satzes 15 gehört jede Idealklasse aus $\mathfrak{B}_\mathfrak{n}$ zu Idealen genau einer Klasse aus $\mathfrak{B}_\mathfrak{N}$, nämlich der Erweiterungsidealklasse.*

2. *Die durch die Ideale aus $\mathfrak{G}(\mathfrak{n}; \mathfrak{N})$ vertretenen Klassen bilden eine Untergruppe G der multiplikativen Gruppe H der Idealklassen aus $\mathfrak{B}_\mathfrak{n}$. Die im Sinne des Satzes 15 zu den Idealen einer gegebenen Klasse aus $\mathfrak{B}_\mathfrak{N}$ gehörigen Idealklassen aus $\mathfrak{B}_\mathfrak{n}$ bilden eine Restklasse von H nach G; die Menge dieser Idealklassen hat also die Mächtigkeit $g(\mathfrak{n}; \mathfrak{N})$.*

Beweis. 1. Ist $\mathfrak{A}$ ein zu $\mathfrak{f}$ teilerfremdes Ideal aus $\mathfrak{B}_\mathfrak{N}$, dessen Verengungsideal in der Klasse a aus $\mathfrak{B}_\mathfrak{n}$ liegt, so ist nach Satz 11

$$\mathfrak{A} = (\mathfrak{A} \cap \mathfrak{n})\mathfrak{N},$$

so daß $\mathfrak{A}$ der Erweiterungsidealklasse von a angehört. Ein solches $\mathfrak{A}$ gibt es aber zu gegebenem a nach Satz 15 wirklich.

2. Die durch 1. gegebene Abbildung der Gruppe H auf die Gruppe H' der Idealklassen aus $\mathfrak{B}_\mathfrak{N}$ erschöpft diese gänzlich, da in jeder Klasse aus $\mathfrak{B}_\mathfrak{N}$ nach Satz 10 ein zu $\mathfrak{f}$ teilerfremdes Ideal liegt. Die Erweiterungsidealklasse des Produktes zweier Idealklassen ist das Produkt der Erweiterungsidealklassen; die Abbildung ist also ein Homomorphismus. Diejenigen Elemente von H, deren Bild das Einheitselement von H' ist, bilden also eine Unter-

gruppe, und zu jedem Element von H' gehören genau die Elemente einer Restklasse von H nach dieser Untergruppe. Nach Definition 3 und Satz 10 besteht diese Untergruppe gerade aus den durch die Elemente von $\mathfrak{G}(\mathfrak{n}; \mathfrak{N})$ vertretenen Idealklassen.

Ist insbesondere $\mathfrak{R}$ der Ring der ganzen rationalen Zahlen, K ein endlicher algebraischer Zahlkörper und $\mathfrak{N}$ die Hauptordnung von K, so kennt man nach Dedekind[25]) das Verhältnis der (endlichen) Anzahl derjenigen Idealklassen in $\mathfrak{n}$, die durch die zu $\mathfrak{f}$ teilerfremden Ideale vertreten werden, zu der (endlichen) Anzahl aller Idealklassen in $\mathfrak{N}$. Da aber $\mathfrak{N}$ Hauptordnung ist und somit jedes Ideal in $\mathfrak{N}$ zugleich zu $\mathfrak{B}_{\mathfrak{N}}$ gehört, so sind nach Satz 12, 14 und 10 die genannten Klassen gerade die sämtlichen Idealklassen aus $\mathfrak{B}_{\mathfrak{n}}$ und $\mathfrak{B}_{\mathfrak{N}}$. Dedekind hat also in Wahrheit auch das Verhältnis der Anzahlen der Idealklassen aus $\mathfrak{B}_{\mathfrak{n}}$ und $\mathfrak{B}_{\mathfrak{N}}$ oder, was nach Satz 17 dasselbe ist, die (hier endliche) Mächtigkeit $g(\mathfrak{n}; \mathfrak{N})$ bestimmt. Eine einfache Division ergibt alsdann das Verhältnis der Anzahlen der Idealklassen aus $\mathfrak{B}_{\mathfrak{n}_1}$ und $\mathfrak{B}_{\mathfrak{n}_2}$, wenn $\mathfrak{n}_1$ und $\mathfrak{n}_2$ zwei beliebige Ordnungen vom Range der Hauptordnung[26]) sind.

§ 4.

Auflösung der linearen Gleichung.

Wegen $\mathfrak{f} \neq (0)$ haben $\mathfrak{n}$ und $\mathfrak{N}$ denselben Unterkörper von K zum Quotientenkörper[27]); er werde in diesem Paragraphen mit T bezeichnet.

Definition 5. *Der Index*[28]) *der Gruppe der gemeinsamen Einheiten von* $\mathfrak{n}$ *und* $\mathfrak{N}$ [*mit positiver Norm*] *in der Obergruppe der sämtlichen Einheiten von* $\mathfrak{N}$ [*mit positiver Norm*] *wird mit* $e(\mathfrak{n}; \mathfrak{N})$ *bezeichnet.*

Definition 6. *Ist* $\mathfrak{o}$ *eine Ordnung, so heißt ein Modul* $\mathfrak{m}$ *gebrochenes Ideal von* $\mathfrak{o}$, *wenn es ein von Null verschiedenes Element* α *von* $\mathfrak{o}$ *gibt, für welches* $\alpha\mathfrak{m}$ *Ideal in* $\mathfrak{o}$ *ist.*

Jedes gebrochene Ideal einer Ordnung $\mathfrak{o}$ liegt ersichtlich im Quotientenkörper von $\mathfrak{o}$.

---

[25]) D. 2, § 10. Man beachte das unten zu Beginn des § 4 Gesagte.

[26]) Nach einem Satz aus G. (S. 548 f.) ist nämlich für zwei solche Ordnungen der Führer $\mathfrak{f}$ wegen der Gleichheit der Quotientenkörper vom Nullideal verschieden. Die bei Grell vorliegende Voraussetzung, daß $\mathfrak{N}$ von $\mathfrak{R}$ verschieden ist, übrigens auch die (oben erfüllte) Voraussetzung, daß $\mathfrak{R}$ Multiplikationsring ist, wird beim Beweis des betreffenden Teils der Behauptung nicht benutzt.

[27]) Nach dem Satz aus G., S. 548 f. Die dortigen Voraussetzungen, daß $\mathfrak{N}$ endlich und von $\mathfrak{R}$ verschieden und $\mathfrak{R}$ Multiplikationsring ist, werden in dem betreffenden Teil des Beweises nicht benutzt.

[28]) Als Index wird die Mächtigkeit der Menge der Restklassen bezeichnet.

Wie üblich, erklärt man eine *Klasse gebrochener Ideale* einer Ordnung $\mathfrak{o}$ als die Gesamtheit aller gebrochenen Ideale $\mathfrak{b}$ von $\mathfrak{o}$, die zu einem festen gebrochenen Ideal $\mathfrak{a}$ von $\mathfrak{o}$ in einer Beziehung von der Gestalt

$$\eta_1 \mathfrak{a} = \eta_2 \mathfrak{b}$$

stehen, wo η_1 und η_2 von Null verschiedene Elemente von $\mathfrak{o}$ sind [und $\frac{\eta_1}{\eta_2}$ positive Norm hat]. Alle gebrochenen Ideale einer Klasse haben dieselbe Ordnung; sie sind umkehrbar, sobald eins von ihnen es ist. Nun ist jede Klasse von Idealen aus $\mathfrak{B}_\mathfrak{o}$ offenbar Teilmenge einer Klasse umkehrbarer gebrochener Ideale von $\mathfrak{o}$, die die Ordnung $\mathfrak{o}$ haben, und nur einer solchen Klasse, und jede Klasse umkehrbarer gebrochener Ideale von $\mathfrak{o}$ mit der Ordnung $\mathfrak{o}$ enthält die Ideale genau einer Idealklasse aus $\mathfrak{B}_\mathfrak{o}$. In diesem Sinne kann man von *zueinander gehörigen* Klassen „ganzer" und gebrochener Ideale reden. Diese Bezeichnung wird im folgenden benutzt.

Satz 18. *Ist* $\mathfrak{Q}$ *ein umkehrbares gebrochenes Ideal von* $\mathfrak{N}$ *und* $\mathfrak{Q}^0 = \mathfrak{N}$, *so hat die Menge der umkehrbaren gebrochenen Ideale* $\mathfrak{m}$ *von* $\mathfrak{n}$ *mit* $\mathfrak{m}^0 = \mathfrak{n}$ *und*

$$\mathfrak{m}\mathfrak{N} = \mathfrak{Q} \tag{2}$$

die Mächtigkeit $g(\mathfrak{n}; \mathfrak{N})\, e(\mathfrak{n}; \mathfrak{N})$.

Im Wortlaut des Satzes kann übrigens nach einem Satz von Krull[29]) der Begriff „gebrochenes Ideal von $\mathfrak{N}$" mit Rücksicht auf die Nebenbedingungen durch den einfacheren Begriff „Modul aus T" ersetzt werden. Denn jeder in T gelegene umkehrbare Modul von der Ordnung $\mathfrak{N}$ hat nach jenem Satz eine endliche $\mathfrak{N}$-Modul-Basis, ergibt also durch Multiplikation mit einem passenden nichtverschwindenden Element von $\mathfrak{N}$ ein Ideal in $\mathfrak{N}$ und ist somit gebrochenes Ideal von $\mathfrak{N}$. — Eine entsprechende Vereinfachung kann man darauf mit $\mathfrak{n}$ statt $\mathfrak{N}$ vornehmen.

Beweis. Zu der $\mathfrak{Q}$ enthaltenden Klasse gebrochener Ideale von $\mathfrak{N}$ gehört eine Idealklasse Q aus $\mathfrak{B}_\mathfrak{N}$. Ist Q die Erweiterungsidealklasse einer Idealklasse k aus $\mathfrak{B}_\mathfrak{n}$, so gibt es in der zu k gehörigen Klasse m gebrochener Ideale von $\mathfrak{n}$ offenbar mindestens ein $\mathfrak{m}$ mit (2). Haben die Moduln $\mathfrak{m}_1$ und $\mathfrak{m}_2$ aus m diese Eigenschaft und wird

$$\mathfrak{m}_1 = \varepsilon \mathfrak{m}_2$$

gesetzt, wo ε ein Element von T [mit positiver Norm] bedeutet, so ergibt sich durch Multiplikation mit $\mathfrak{Q}^{-1}$, daß

$$\mathfrak{N} = \varepsilon \mathfrak{N},$$

---

[29]) Ein Hauptsatz über umkehrbare Ideale (Math. Zeitschr. 31 (1930), S. 558).

also ε eine Einheit von $\mathfrak{N}$ ist. Umgekehrt, wenn ε eine Einheit von $\mathfrak{N}$ [mit positiver Norm] ist, so ist mit $\mathfrak{m}$ auch $\varepsilon\mathfrak{m}$ ein Element von m mit der Eigenschaft (2); sind ε_1 und ε_2 Einheiten von $\mathfrak{N}$ [mit positiver Norm], so ist dann und nur dann

$$\varepsilon_1 \mathfrak{m} = \varepsilon_2 \mathfrak{m},$$

wenn

$$\frac{\varepsilon_1}{\varepsilon_2} \mathfrak{m} = \mathfrak{m}$$

ist, d. h. wenn

$$\frac{\varepsilon_1}{\varepsilon_2} \mathfrak{n} = \mathfrak{n},$$

d. h. $\frac{\varepsilon_1}{\varepsilon_2}$ eine Einheit von $\mathfrak{n}$ ist. Die Menge der in m enthaltenen gebrochenen Ideale $\mathfrak{m}$ mit der Eigenschaft (2) hat also die Mächtigkeit $e(\mathfrak{n}; \mathfrak{N})$.

Ist umgekehrt $\mathfrak{m}$ ein gebrochenes Ideal von $\mathfrak{n}$ von der im Satze verlangten Art und bedeutet k die zur Klasse von $\mathfrak{m}$ gehörige Idealklasse aus $\mathfrak{B}_\mathfrak{n}$, so ist Q die Erweiterungsidealklasse von k.

Die Menge der Lösungen von (2) mit den geforderten Eigenschaften hat also nach Satz 17 die Mächtigkeit $g(\mathfrak{n}; \mathfrak{N})\, e(\mathfrak{n}; \mathfrak{N})$.

Satz 19. *Es liege eine Ordnung* $\mathfrak{n}_1$ *mit*

$$\mathfrak{n}\,\mathfrak{n}_1 = \mathfrak{N}$$

vor. Es sei $\mathfrak{M}$ *ein umkehrbares gebrochenes Ideal von* $\mathfrak{N}$ *mit der Ordnung* $\mathfrak{N}$, *ebenso* $\mathfrak{m}_1$ *ein umkehrbares gebrochenes Ideal von* $\mathfrak{n}_1$ *mit der Ordnung* $\mathfrak{n}_1$. *Dann hat die Menge der umkehrbaren gebrochenen Ideale* $\mathfrak{m}$ *von* $\mathfrak{n}$ *mit* $\mathfrak{m}^0 = \mathfrak{n}$ *und*

$$\mathfrak{m}\,\mathfrak{m}_1 = \mathfrak{M} \tag{3}$$

die Mächtigkeit $g(\mathfrak{n}; \mathfrak{N})\, e(\mathfrak{n}; \mathfrak{N})$.

Statt „gebrochenes Ideal von $\mathfrak{N}$ ($\mathfrak{n}$)“ kann in Anbetracht der Nebenbedingungen wiederum „Modul aus T“, statt „gebrochenes Ideal von $\mathfrak{n}_1$“ in Anbetracht der Nebenbedingungen „Modul aus dem Quotientenkörper von $\mathfrak{n}_1$“ gesagt werden.

Beweis. Für einen beliebigen Modul $\mathfrak{m}$ mit $\mathfrak{m}^0 \geq \mathfrak{n}$ besteht dann und nur dann die Gleichung (3), wenn

$$\mathfrak{m}\,\mathfrak{N} = \mathfrak{M}\,\mathfrak{m}_1^{-1} \tag{4}$$

ist. Denn einerseits folgt aus (3)

$$\mathfrak{m}\,\mathfrak{N} = \mathfrak{m}\,\mathfrak{n}\,\mathfrak{n}_1 = \mathfrak{m}\,\mathfrak{n}_1 = \mathfrak{m}\,\mathfrak{m}_1\,\mathfrak{m}_1^{-1} = \mathfrak{M}\,\mathfrak{m}_1^{-1};$$

andererseits folgt aus (4)

$$\mathfrak{m}\,\mathfrak{m}_1 = \mathfrak{m}\,\mathfrak{n}\,\mathfrak{n}_1\,\mathfrak{m}_1 = \mathfrak{m}\,\mathfrak{N}\,\mathfrak{m}_1 = \mathfrak{M}\,\mathfrak{m}_1^{-1}\,\mathfrak{m}_1 = \mathfrak{M}\,\mathfrak{n}_1 = \mathfrak{M}.$$

Damit ist alles auf die Lösung der Gleichung (4) zurückgeführt. Nach Satz 2 und Satz 3 ist der Modul $\mathfrak{Q} = \mathfrak{M}\mathfrak{m}_1^{-1}$ umkehrbar und $\mathfrak{Q}^0 = \mathfrak{N}\mathfrak{n}_1 = \mathfrak{N}$. Nach Definition 6 gibt es ein nichtverschwindendes Element α von $\mathfrak{n}_1$, für welches $\alpha\mathfrak{m}_1 \leqq \mathfrak{n}_1$ ist. Ist β ein beliebiges von Null verschiedenes Element von $\mathfrak{m}_1$, so ist $\beta\mathfrak{m}_1^{-1} \leqq \mathfrak{n}_1$, also $(\alpha\beta)\mathfrak{m}_1^{-1} \leqq \mathfrak{n}_1$ und hierin $\alpha\beta \in \mathfrak{n}_1$. Also liegt $\mathfrak{m}_1^{-1}$ und somit auch $\mathfrak{Q}$ in T. Ist ferner γ ein von Null verschiedenes Element von $\mathfrak{N}$, für welches $\gamma\mathfrak{M} \leqq \mathfrak{N}$ wird, so ist

$$(\alpha\beta\gamma)\mathfrak{Q} \leqq \mathfrak{N}\mathfrak{n}_1 = \mathfrak{N},$$

hierbei $\alpha\beta\gamma \in \mathfrak{N}$. Der Modul $\mathfrak{Q}$ erfüllt also die Voraussetzung des Satzes 18. Aus diesem folgt sofort die Behauptung.

§ 5.

Normen.

Ein Modul (insbesondere eine Ordnung) heißt *endlich*, wenn er eine endliche $\mathfrak{R}$-Modul-Basis hat.

Ist $\mathfrak{R}$ Hauptidealring, so gestattet nach einer von Grell[30]) herrührenden Verallgemeinerung eines Gruppensatzes von Prüfer[31]) jeder endliche, vom Nullmodul verschiedene Modul $\mathfrak{m}$ eine Darstellung als direkte Summe

$$\mathfrak{m} = \alpha_1\mathfrak{R} + \cdots + \alpha_n\mathfrak{R} \quad (n \geqq 1;\ \alpha_i \in \mathfrak{m},\ \alpha_i \neq 0 \text{ für } i = 1, \ldots, n); \tag{5}$$

die Gesamtheit der Moduln $\alpha_1\mathfrak{R}, \ldots, \alpha_n\mathfrak{R}$ ist dabei bis auf $\mathfrak{R}$-Modul-Isomorphie eindeutig bestimmt. Eine direkte Summendarstellung von $\mathfrak{m}$ in der Gestalt (5) wird im folgenden kurz eine *Normaldarstellung* genannt; die durch den Modul $\mathfrak{m}$ eindeutig bestimmte Zahl n heißt der *Rang* von $\mathfrak{m}$ in bezug auf $\mathfrak{R}$. Sind $\mathfrak{a}$ und $\mathfrak{b}$ endliche, vom Nullmodul verschiedene Moduln von gleichem Rang n und ist $\mathfrak{a} \geqq \mathfrak{b}$, so gibt es nach dem Hauptsatz über die Elementarteiler einer Matrix[32]) zwei Normaldarstellungen

$$\mathfrak{a} = \alpha_1\mathfrak{R} + \cdots + \alpha_n\mathfrak{R},$$
$$\mathfrak{b} = \beta_1\mathfrak{R} + \cdots + \beta_n\mathfrak{R},$$

bei denen jedes $\frac{\beta_i}{\alpha_i}$ $(i = 1, \ldots, n)$ Element von $\mathfrak{R}$ und für $i < n$ sogar Element von $\frac{\beta_{i+1}}{\alpha_{i+1}}\mathfrak{R}$ ist[33]). Da die Elementarteiler einer Matrix bis auf Einheitsfaktoren eindeutig bestimmt sind und überdies zu isomorphen Rest-

[30]) G., S. 532.

[31]) Theorie der Abelschen Gruppen, I. Grundeigenschaften (Math. Zeitschr. 20 (1924), S. 165–187), § 2.

[32]) Vgl. etwa van der Waerden, Moderne Algebra 2 (1931), § 106.

[33]) Vgl. auch G., S. 533.

klassen-$\mathfrak{R}$-Moduln dieselben Elementarteiler gehören[34]), so sind die $\frac{\beta_i}{\alpha_i}$ schon durch den abstrakten Restklassen-$\mathfrak{R}$-Modul von $\mathfrak{a}$ nach $\mathfrak{b}$ bis auf Einheitsfaktoren aus $\mathfrak{R}$ eindeutig bestimmt.

Von nun an sei $\mathfrak{R}$ in diesem und dem folgenden Paragraphen speziell Hauptidealring.

Definition 7. *Die vom Nullmodul verschiedenen endlichen Moduln* $\mathfrak{a}$ *und* $\mathfrak{b}$ *seien vom selben Rang* n. *Zwei Normaldarstellungen seien*

$$\mathfrak{a} = \alpha_1 \mathfrak{R} + \cdots + \alpha_n \mathfrak{R},$$
$$\mathfrak{b} = \beta_1 \mathfrak{R} + \cdots + \beta_n \mathfrak{R}.$$

Es sei $\mathfrak{a} \geqq \mathfrak{b}$, *also für* $k = 1, \ldots, n$ *etwa*

$$\beta_k = \sum_{i=1}^{n} \alpha_i r_{ik}, \quad r_{ik} \in \mathfrak{R} \qquad (i = 1, \ldots, n).$$

Dann heißt die Matrix (r_{ik}) (i *Zeilen-,* k *Spaltenindex*) *eine Übergangsmatrix von* $\mathfrak{a}$ *nach* $\mathfrak{b}$. *Die Gesamtheit aller Übergangsmatrizes von* $\mathfrak{a}$ *nach* $\mathfrak{b}$ *wird mit* $\mathsf{N}(\mathfrak{a}|\mathfrak{b})$ *bezeichnet.*

Man erhält $\mathsf{N}(\mathfrak{a}|\mathfrak{b})$ offenbar aus einer beliebigen Übergangsmatrix von $\mathfrak{a}$ nach $\mathfrak{b}$, indem man diese beiderseitig mit allen n-reihigen Matrizen multipliziert, deren Elemente zu $\mathfrak{R}$ gehören und deren Determinanten Einheiten von $\mathfrak{R}$ sind. Diese Matrizen sind gerade die Einheiten des Ringes Λ aller n-reihigen Matrizen aus Elementen von $\mathfrak{R}$.

Satz 20 [33]). *Es seien* $\mathfrak{a}_1, \mathfrak{b}_1, \mathfrak{a}_2, \mathfrak{b}_2$ *endliche, vom Nullmodul verschiedene Moduln vom selben Rang; es sei* $\mathfrak{a}_1 \geqq \mathfrak{b}_1$, $\mathfrak{a}_2 \geqq \mathfrak{b}_2$. *Die Restklassenmoduln von* $\mathfrak{a}_1$ *nach* $\mathfrak{b}_1$ *und von* $\mathfrak{a}_2$ *nach* $\mathfrak{b}_2$ *seien als* $\mathfrak{R}$-*Moduln isomorph. Dann ist*

$$\mathsf{N}(\mathfrak{a}_1|\mathfrak{b}_1) = \mathsf{N}(\mathfrak{a}_2|\mathfrak{b}_2).$$

Beweis. Nach den obigen Bemerkungen können vier Normaldarstellungen

$$\left.\begin{aligned} \mathfrak{a}_i &= \alpha_{i1} \mathfrak{R} + \cdots + \alpha_{in} \mathfrak{R}, \\ \mathfrak{b}_i &= \beta_{i1} \mathfrak{R} + \cdots + \beta_{in} \mathfrak{R} \end{aligned}\right\} \qquad (i = 1, 2)$$

und Elemente $d_1, \ldots, d_n$ von $\mathfrak{R}$ so gefunden werden, daß dabei

$$\beta_{ik} = d_k \alpha_{ik} \qquad (i = 1, 2;\ k = 1, \ldots, n)$$

ist. Die Matrix $\begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix}$ ist eine Übergangsmatrix von $\mathfrak{a}_1$ nach $\mathfrak{b}_1$ und zugleich von $\mathfrak{a}_2$ nach $\mathfrak{b}_2$; daraus folgt sofort die Behauptung.

---

[34]) Vgl. van der Waerden, a. a. O., § 107.

Satz 21. *Sind* $\mathfrak{a}, \mathfrak{b}, \mathfrak{c}$ *endliche, vom Nullmodul verschiedene Moduln von gleichem Rang und ist* $\mathfrak{a} \geqq \mathfrak{b} \geqq \mathfrak{c}$, *sind ferner* M_1, M_2, M_3 *Übergangsmatrizes von* $\mathfrak{a}$ *nach* $\mathfrak{b}$ *bzw. von* $\mathfrak{b}$ *nach* $\mathfrak{c}$ *bzw. von* $\mathfrak{a}$ *nach* $\mathfrak{c}$, *so gibt es in* Λ *drei Einheiten* P_1, P_2, P_3 *mit*

$$P_1 M_1 P_2 M_2 P_3 = M_3 .$$

Beweis. Für $\mathfrak{a}, \mathfrak{b}, \mathfrak{c}$ hat man je ein Paar von Normaldarstellungen mittels gewisser Basen $(\alpha), (\alpha')$ bzw. $(\beta), (\beta')$ bzw. $(\gamma), (\gamma')$ von gemeinsamer Länge n derart, daß die Matrizengleichungen

$$\begin{aligned}(\beta_1, \ldots, \beta_n) &= (\alpha_1', \ldots, \alpha_n') M_1 ,\\ (\gamma_1, \ldots, \gamma_n) &= (\beta_1', \ldots, \beta_n') M_2 ,\\ (\gamma_1', \ldots, \gamma_n') &= (\alpha_1, \ldots, \alpha_n) M_3\end{aligned}$$

bestehen. In Λ gibt es andererseits Einheiten P_1, P_2, P_3 mit

$$\begin{aligned}(\alpha_1', \ldots, \alpha_n') &= (\alpha_1, \ldots, \alpha_n) P_1 ,\\ (\beta_1', \ldots, \beta_n') &= (\beta_1, \ldots, \beta_n) P_2 ,\\ (\gamma_1', \ldots, \gamma_n') &= (\gamma_1, \ldots, \gamma_n) P_3 .\end{aligned}$$

Daraus folgt der Satz mit Rücksicht auf die lineare Unabhängigkeit von $\alpha_1, \ldots, \alpha_n$ in bezug auf den Quotientenkörper von $\mathfrak{R}$.

Satz 22. *Ist* $\mathfrak{n}$ *eine endliche Ordnung und sind* $\mathfrak{a}$ *und* $\mathfrak{b}$ *zwei vom Nullideal verschiedene Ideale in* $\mathfrak{n}$, *so ist*

$$\mathsf{N}(\mathfrak{a} + \mathfrak{b} | \mathfrak{a}) = \mathsf{N}(\mathfrak{b} | \mathfrak{a} \cap \mathfrak{b}).$$

Beweis. Ist

$$\mathfrak{n} = \omega_1 \mathfrak{R} + \cdots + \omega_n \mathfrak{R}$$

eine Normaldarstellung von $\mathfrak{n}$, $\mathfrak{c}$ ein Ideal in $\mathfrak{n}$, γ ein von Null verschiedenes Element von $\mathfrak{c}$, so sind die in $\mathfrak{c}$ gelegenen Größen

$$\gamma\omega_1, \ldots, \gamma\omega_n$$

linear-unabhängig in bezug auf den Quotientenkörper von $\mathfrak{R}$. Da überdies $\mathfrak{R}$ Hauptidealring ist, so findet man in bekannter Weise[35]) n Größen

$$\left.\begin{aligned}\gamma_1 &= r_{11}\omega_1 ,\\ \gamma_2 &= r_{21}\omega_1 + r_{22}\omega_2 ,\\ &\cdots\cdots\cdots ,\\ \gamma_n &= r_{n1}\omega_1 + r_{n2}\omega_2 + \cdots + r_{nn}\omega_n\end{aligned}\right\} \quad (r_{ik} \in \mathfrak{R} \text{ für } i, k = 1, \ldots, n \text{ und } i \geqq k),$$

mit deren Hilfe $\mathfrak{c}$ die Normaldarstellung

$$\mathfrak{c} = \gamma_1 \mathfrak{R} + \cdots + \gamma_n \mathfrak{R}$$

[35]) Vgl. etwa H. Weber, Lehrbuch der Algebra, 2. Aufl., 2 (1899), S. 602f.

zuläßt. Jedes vom Nullideal verschiedene Ideal in $\mathfrak{n}$ ist somit ein endlicher Modul und hat mit $\mathfrak{n}$ gleichen Rang. Mit $\mathfrak{a}$ und $\mathfrak{b}$ ist $\mathfrak{a}+\mathfrak{b}$ und (wegen $\mathfrak{a}\cap\mathfrak{b} \geqq \mathfrak{ab} \neq (0)$) auch $\mathfrak{a}\cap\mathfrak{b}$ ein vom Nullideal verschiedenes Ideal in $\mathfrak{n}$; zudem ist $\mathfrak{a}+\mathfrak{b} \geqq \mathfrak{a}$, $\mathfrak{b} \geqq \mathfrak{a}\cap\mathfrak{b}$. Bekanntlich[36]) sind nun die Restklassenmoduln von $\mathfrak{a}+\mathfrak{b}$ nach $\mathfrak{a}$ und von $\mathfrak{b}$ nach $\mathfrak{a}\cap\mathfrak{b}$ als $\mathfrak{R}$-Moduln isomorph. Aus Satz 20 folgt also die Behauptung.

Satz 23. *Es sei $\mathfrak{n}$ eine endliche Ordnung, $\mathfrak{a}$ ein Ideal aus dem Bereich $\mathfrak{B}_\mathfrak{n}$, $\mathfrak{b}$ ein beliebiges vom Nullideal verschiedenes Ideal in $\mathfrak{n}$. Sind dann bzw. M_1, M_2, M Übergangsmatrizes von $\mathfrak{n}$ nach $\mathfrak{a}, \mathfrak{b}, \mathfrak{ab}$, so gibt es in Λ drei Einheiten P_1, P_2, P_3 mit*

$$P_1 M_1 P_2 M_2 P_3 = M.$$

Beweis[37]). Wegen $\mathfrak{n} \geqq \mathfrak{a} \geqq \mathfrak{ab}$ braucht nach Satz 21 nur gezeigt zu werden, daß es in Λ zwei Einheiten P_4 und P_5 gibt von der Beschaffenheit, daß $P_4 M_2 P_5$ eine Übergangsmatrix von $\mathfrak{a}$ nach $\mathfrak{ab}$ ist. Denn dann ist nach jenem Satz bei passenden Einheiten P_1, P_6, P_7 aus Λ

$$P_1 M_1 P_6 P_4 M_2 P_5 P_7 = M.$$

Nach Satz 9 gibt es nun in $\mathfrak{n}$ ein von Null verschiedenes Element α mit

$$\mathfrak{ab} + \alpha\mathfrak{n} = \mathfrak{a}. \tag{6}$$

Nach Satz 22 wird

$$\mathsf{N}(\mathfrak{a}|\mathfrak{ab}) = \mathsf{N}(\mathfrak{ab}+\alpha\mathfrak{n}|\mathfrak{ab}) = \mathsf{N}(\alpha\mathfrak{n}|\mathfrak{ab}\cap\alpha\mathfrak{n}). \tag{7}$$

Nun ist nach (6)

$$\mathfrak{a}(\mathfrak{ab}\cap\alpha\mathfrak{n}) = (\mathfrak{ab}+\alpha\mathfrak{n})(\mathfrak{ab}\cap\alpha\mathfrak{n}) \subseteqq \mathfrak{ab}\cdot\alpha\mathfrak{n} = \alpha\mathfrak{ab};$$

Multiplikation mit $\mathfrak{a}^{-1}$ liefert

$$\mathfrak{ab}\cap\alpha\mathfrak{n} \subseteqq \alpha\mathfrak{b}.$$

Nach (7) und Satz 21 gibt es also drei Übergangsmatrizes M_3, M_4, M_5 von $\alpha\mathfrak{n}$ nach $\alpha\mathfrak{b}$ bzw. von $\alpha\mathfrak{b}$ nach $\mathfrak{ab}\cap\alpha\mathfrak{n}$ bzw. von $\mathfrak{a}$ nach $\mathfrak{ab}$ mit

$$M_3 M_4 = M_5.$$

M_3 ist zugleich eine Übergangsmatrix von $\mathfrak{n}$ nach $\mathfrak{b}$, also von der Gestalt $P_1 M_2 P_8$, wo P_4 und P_8 Einheiten von Λ sind; somit ist

$$P_4 M_2 P_8 M_4 = M_5. \tag{8}$$

---

[36]) Vgl. N., S. 40.

[37]) Das Verfahren lehnt sich an den Beweis des Normenmultiplikationssatzes für Hauptordnungen von Zahlkörpern in D. 1, S. 564 an. Das dortige Verfahren leistet hier zunächst weniger; man muß es, wie oben geschehen, zweimal durchführen, um zum Ziele zu kommen.

Nach Satz 6 gibt es ferner in $\mathfrak{B}_\mathfrak{n}$ ein Ideal $\mathfrak{x}$ und in $\mathfrak{n}$ ein von Null verschiedenes Element β mit

$$\mathfrak{a}\mathfrak{x} = \beta\mathfrak{n}.$$

Da das Ideal $\beta\mathfrak{x}$ zu $\mathfrak{B}_\mathfrak{n}$ gehört, so folgt aus Satz 9 das Vorhandensein eines von Null verschiedenen Elements γ von $\mathfrak{n}$ mit

$$\beta\mathfrak{x}\mathfrak{b} + \gamma\mathfrak{n} = \beta\mathfrak{x},$$

woraus sich durch Multiplikation mit $\mathfrak{a}$ ergibt

$$\beta^2\mathfrak{b} + \gamma\mathfrak{a} = \beta^2\mathfrak{n}.$$

Also ist, wiederum nach Satz 22,

$$(9) \quad \mathsf{N}(\mathfrak{n}|\mathfrak{b}) = \mathsf{N}(\beta^2\mathfrak{n}|\beta^2\mathfrak{b}) = \mathsf{N}(\beta^2\mathfrak{b} + \gamma\mathfrak{a}|\beta^2\mathfrak{b}) = \mathsf{N}(\gamma\mathfrak{a}|\beta^2\mathfrak{b} \cap \gamma\mathfrak{a}).$$

Es gilt nun

$$\beta^2(\beta^2\mathfrak{b} \cap \gamma\mathfrak{a}) = (\beta^2\mathfrak{b} + \gamma\mathfrak{a})\,(\beta^2\mathfrak{b} \cap \gamma\mathfrak{a}) \leqq \beta^2\gamma\mathfrak{a}\mathfrak{b}, \quad \beta^2\mathfrak{b} \cap \gamma\mathfrak{a} \leqq \gamma\mathfrak{a}\mathfrak{b}.$$

Nach (9) und Satz 21 gibt es also zwei Übergangsmatrizes M_6, M_7 von $\gamma\mathfrak{a}$ nach $\gamma\mathfrak{a}\mathfrak{b}$ bzw. von $\gamma\mathfrak{a}\mathfrak{b}$ nach $\beta^2\mathfrak{b} \cap \gamma\mathfrak{a}$ mit

$$M_6 M_7 = M_2.$$

Da M_6 die Gestalt $P_9 M_5 P_{10}$ hat, wo P_9 und P_{10} Einheiten von Λ sind, so folgt

$$P_9 M_5 P_{10} M_7 = M_2,$$

also nach (8)

$$P_9 P_4 M_2 P_8 M_4 P_{10} M_7 = M_2.$$

Durch Übergang zu den Determinanten sieht man, daß die Determinante von M_4 eine Einheit von $\mathfrak{R}$, demnach M_4 eine Einheit von Λ ist. Nach (8) ist damit der Satz bewiesen.

Definition 8. *Der vom Nullmodul verschiedene Modul* $\mathfrak{a}$ *sei ein Ring* (also Ideal in seiner Ordnung); *die Ordnung* $\mathfrak{n}$ *von* $\mathfrak{a}$ *sei endlich. Das aus der Determinante einer beliebigen Übergangsmatrix von* $\mathfrak{n}$ *nach* $\mathfrak{a}$ *abgeleitete Ideal in* $\mathfrak{R}$ *heißt dann die Norm von* $\mathfrak{a}$ *und wird mit* $N(\mathfrak{a})$ *bezeichnet.*

Dieses Ideal ist von der speziellen Wahl der Übergangsmatrix unabhängig, da die Determinante der letzteren bis auf einen Einheitsfaktor aus $\mathfrak{R}$ bestimmt ist.

Satz 24. *Ist* $\mathfrak{n}$ *eine endliche Ordnung,* $\mathfrak{a}$ *und* $\mathfrak{b}$ *zwei vom Nullideal verschiedene Ideale in* $\mathfrak{n}$, *die beide die Ordnung* $\mathfrak{n}$ *haben und von denen mindestens eins umkehrbar ist, so ist*

$$N(\mathfrak{a})\,N(\mathfrak{b}) = N(\mathfrak{a}\mathfrak{b}).$$

Beweis. Ohne Beschränkung der Allgemeinheit kann man annehmen, daß $\mathfrak{a}$ umkehrbar ist, also in $\mathfrak{B}_n$ liegt. Da aus $\alpha \in \mathsf{K}$, $\alpha\mathfrak{a}\mathfrak{b} \leqq \mathfrak{a}\mathfrak{b}$ durch Multiplikation mit $\mathfrak{a}^{-1}$ folgt $\alpha\mathfrak{b} \leqq \mathfrak{b}$, $\alpha \in \mathfrak{n}$, so hat $\mathfrak{a}\mathfrak{b}$ die Ordnung $\mathfrak{n}$. Die Behauptung ergibt sich dann aus Satz 23 durch Übergang zu den Determinanten.

§ 6.

Umkehrbare Ideale in Erweiterungskörpern erster Art eines Hauptidealrings. Normenrelationen.

Satz 25. *Ist s Element eines im Quotientenkörper ganz-abgeschlossenen*[38]) *(nullteilerfreien) Ringes $\mathfrak{S}$ mit Einheitselement und ist $\mathfrak{T}$ eine nur aus $\mathfrak{S}$-ganzen Größen bestehende $\mathfrak{S}$-Ordnung*[13]) *in einem Oberring von $\mathfrak{S}$, so ist*

$$s\mathfrak{T} \cap \mathfrak{S} = s\mathfrak{S}.$$

Beweis. Im Falle $s = 0$ ist der Satz klar; es sei also $s \neq 0$. Es wird

$$s\mathfrak{S} \leqq s\mathfrak{T},$$

$$s\mathfrak{S} \leqq \mathfrak{S},$$

also

$$s\mathfrak{S} \leqq s\mathfrak{T} \cap \mathfrak{S}.$$

Andererseits hat jedes Element von $s\mathfrak{T} \cap \mathfrak{S}$ die Gestalt st, wo t in $\mathfrak{T}$ liegt, und ist zugleich Element von $\mathfrak{S}$; t liegt im Quotientenkörper von $\mathfrak{S}$, als $\mathfrak{S}$-ganze Größe also in $\mathfrak{S}$ selbst, so daß

$$st \in s\mathfrak{S},$$

$$s\mathfrak{T} \cap \mathfrak{S} \leqq s\mathfrak{S}$$

ist, woraus nach dem Obigen die Behauptung folgt.

Zu den nur aus $\mathfrak{S}$-ganzen Größen bestehenden $\mathfrak{S}$-Ordnungen gehören insbesondere die endlichen $\mathfrak{S}$-Ordnungen[39]).

Von nun an bedeute T einen in K gelegenen endlichen Erweiterungskörper erster Art des Quotientenkörpers P von $\mathfrak{R}$. Der Grad von T in bezug auf P wird mit n bezeichnet. K enthalte einen (also auch nur einen) zu T gehörigen in bezug auf P Galoisschen Körper. Obere Indizes $1, \ldots, n$ beziehen sich auf den Übergang zu den n zu T in bezug auf P konjugierten Körpern in diesem Galoisschen Erweiterungskörper von P.

---

[38]) Zu diesem Begriff s. N., S. 33. Ob man hier und im folgenden den alten Ganzheitsbegriff von Dedekind (D. 1, § 173; allgemein gefaßt: N., § 1) oder den neueren von van der Waerden (a. a. O., § 98) zugrunde legt, ist für das Bestehen des Satzes gleichgültig.

[39]) Das folgt aus N., S. 32 oder ist selbstverständlich, je nach dem Ganzheitsbegriff.

Satz 26. *Es sei* $\mathfrak{n}$ *eine endliche Ordnung vom Quotientenkörper* T, *und* γ *sei ein von Null verschiedenes Element von* $\mathfrak{n}$. *Dann ist*

$$N(\gamma\mathfrak{n}) = \gamma^{(1)} \cdots \gamma^{(n)} \mathfrak{R}.$$ [40])

Beweis. $\mathfrak{n}$ hat den Rang n. Es sei

$$\mathfrak{n} = \alpha_1 \mathfrak{R} + \cdots + \alpha_n \mathfrak{R}$$

eine Normaldarstellung. Dann ist eine Normaldarstellung des Ideals $\gamma\mathfrak{n}$ durch

$$\gamma\mathfrak{n} = (\gamma\alpha_1)\mathfrak{R} + \cdots + (\gamma\alpha_n)\mathfrak{R}$$

gegeben, und für $k = 1, \ldots, n$ ist etwa

$$\gamma\alpha_k = \sum_{i=1}^{n} \alpha_i r_{ik}, \qquad r_{ik} \in \mathfrak{R} \qquad (i = 1, \ldots, n).$$

Da die α_i linear-unabhängig in bezug auf P sind, so ist die Matrix (r_{ik}) gerade diejenige, die dem Element γ bei der Darstellung[41]) des Körpers T im Körper P mittels der Basis $(\alpha_1, \ldots, \alpha_n)$ des Darstellungsmoduls T entspricht. Nun ist von E. Noether[42]) der folgende Satz bewiesen: Ist T Erweiterungskörper erster Art eines Körpers P und vom Grade n über P und ist Ω der zu T gehörige in bezug auf P Galoissche Körper, so ist die „Norm" eines Elementes γ von T gleich dem Produkte der n Konjugierten von γ in Ω; wobei aber die „Norm" nicht ein Hauptideal in einem Grundring bedeutet, sondern das $(-1)^n$fache der gemeinsamen Determinante derjenigen Matrizes ist, die dem γ bei den Darstellungen von T in P mittels der verschiedenen Basen des Darstellungsmoduls T entsprechen. Aus diesem Satz ergibt sich, daß im vorliegenden Falle die Matrix (r_{ik}) die Determinante $(-1)^n \gamma^{(1)} \cdots \gamma^{(n)}$ hat. Daraus folgt

$$N(\gamma\mathfrak{n}) = |r_{ik}| \mathfrak{R} = \gamma^{(1)} \cdots \gamma^{(n)} \mathfrak{R}.$$

Satz 27. *Es sei* q *ein von Null verschiedenes Element von* $\mathfrak{R}$, $\mathfrak{n}$ *eine endliche Ordnung vom Quotientenkörper* T, $\mathfrak{a}$ *ein zu* $q\mathfrak{n}$ *teilerfremdes Ideal aus dem Bereich* $\mathfrak{B}_\mathfrak{n}$. *Dann ist in* $\mathfrak{R}$ *das Ideal* $N(\mathfrak{a})$ *zu* $q\mathfrak{R}$ *teilerfremd.*

---

[40]) Aus Satz 26 folgt nebenbei der Satz: Es seien $\mathfrak{n}$ und $\mathfrak{n}_1$ endliche Ordnungen vom selben Quotientenkörper; dieser sei von erster Art über P; es sei $\mathfrak{n} \subseteq \mathfrak{n}_1$; es sei γ ein von Null verschiedenes Element von $\mathfrak{n}$; dann ist $N(\gamma\mathfrak{n}) = N(\gamma\mathfrak{n}_1)$. Dieser Satz läßt sich indessen auch unmittelbar auf Grund der Definition 8 leicht beweisen; dabei kann die Voraussetzung, daß der Quotientenkörper von $\mathfrak{n}$ und $\mathfrak{n}_1$ von erster Art ist, noch gestrichen werden.

[41]) Zu diesem Begriff s. E. Noether, Der Diskriminantensatz für die Ordnungen eines algebraischen Zahl- oder Funktionenkörpers (Journal für die reine und angewandte Mathematik 157 (1927), S. 82–104), § 4.

[42]) In § 6,4. der in Fußnote [41]) erwähnten Arbeit.

Beweis. Zunächst sei $\mathfrak{a}$ ein Hauptideal $\gamma\mathfrak{n}$ ($\gamma \in \mathfrak{n}$). Wird die Ordnung $\mathfrak{n}^{(1)} \cdots \mathfrak{n}^{(n)}$ mit $\mathfrak{N}$ bezeichnet, so folgt aus

$$\gamma\mathfrak{n} + q\mathfrak{n} = \mathfrak{n},$$

daß

$$\gamma\mathfrak{N} + q\mathfrak{N} = \mathfrak{N},$$
$$\gamma^{(i)}\mathfrak{N} + q\mathfrak{N} = \mathfrak{N} \qquad (i = 1, \ldots, n)$$

ist. Innerhalb des Ringes $\mathfrak{N}$ sind also die $\gamma^{(i)}\mathfrak{N}$ und somit auch ihr Produkt $(\gamma^{(1)}\mathfrak{N}) \cdots (\gamma^{(n)}\mathfrak{N}) = \gamma^{(1)} \cdots \gamma^{(n)}\mathfrak{N}$ zu $q\mathfrak{N}$ teilerfremd:

$$\gamma^{(1)} \cdots \gamma^{(n)}\mathfrak{N} + q\mathfrak{N} = \mathfrak{N}. \tag{10}$$

Da $\gamma^{(1)} \cdots \gamma^{(n)}$ in P, also wegen der Ganzheit in $\mathfrak{R}$ liegt und $\mathfrak{R}$ Hauptidealring ist, so gibt es in $\mathfrak{R}$ ein Element r, für welches

$$\gamma^{(1)} \cdots \gamma^{(n)}\mathfrak{R} + q\mathfrak{R} = r\mathfrak{R}$$

ist. Aus (10) folgt

$$r\mathfrak{N} = r\mathfrak{R} \cdot \mathfrak{N} = \mathfrak{N},$$
$$\mathfrak{R} = \mathfrak{N} \cap \mathfrak{R} = r\mathfrak{N} \cap \mathfrak{R}.$$

Nach Satz 25 ist

$$r\mathfrak{N} \cap \mathfrak{R} = r\mathfrak{R},$$

also

$$\mathfrak{R} = r\mathfrak{R},$$
$$\gamma^{(1)} \cdots \gamma^{(n)}\mathfrak{R} + q\mathfrak{R} = \mathfrak{R}.$$

Nach Satz 26 ist

$$\gamma^{(1)} \cdots \gamma^{(n)}\mathfrak{R} = N(\gamma\mathfrak{n}) = N(\mathfrak{a}),$$

also

$$N(\mathfrak{a}) + q\mathfrak{R} = \mathfrak{R}.$$

Im allgemeinen Falle sei nach Satz 8 ein zu $q\mathfrak{n}$ teilerfremdes Ideal $\mathfrak{b}$ aus $\mathfrak{B}_\mathfrak{n}$ gewählt, für welches $\mathfrak{a}\mathfrak{b}$ ein Hauptideal wird. Dann ist nach Satz 24

$$N(\mathfrak{a})\,N(\mathfrak{b}) = N(\mathfrak{a}\mathfrak{b}),$$

also $N(\mathfrak{a})$ Teiler von $N(\mathfrak{a}\mathfrak{b})$. Als Produkt zweier zu $q\mathfrak{n}$ teilerfremden Ideale ist $\mathfrak{a}\mathfrak{b}$ zu $q\mathfrak{n}$ teilerfremd. Nach dem schon erledigten Sonderfall ist also innerhalb $\mathfrak{R}$ das Ideal $N(\mathfrak{a}\mathfrak{b})$ und folglich um so mehr sein Teiler $N(\mathfrak{a})$ zu $q\mathfrak{R}$ teilerfremd.

Satz 28. *Ist $\mathfrak{n}$ eine endliche Ordnung vom Quotientenkörper T, $\mathfrak{a}$ ein Ideal aus $\mathfrak{B}_\mathfrak{n}$, so ist $N(\mathfrak{a})$ der größte gemeinsame Teiler der Ideale $N(\alpha\mathfrak{n})$, wo α die von Null verschiedenen Elemente von $\mathfrak{a}$ durchläuft.*

Beweis. Ist $\alpha \in \mathfrak{a}$, $\alpha \neq 0$, so enthält $\mathfrak{B}_\mathfrak{n}$ nach Satz 5 ein Ideal $\mathfrak{b}$ mit

$$\alpha\mathfrak{n} = \mathfrak{a}\mathfrak{b}.$$

Nach Satz 24 ist

$$N(\mathfrak{a})\,N(\mathfrak{b}) = N(\alpha\mathfrak{n}),$$

also $N(\mathfrak{a})$ gemeinsamer Teiler aller $N(\alpha\mathfrak{n})$. Andrerseits gibt es nach Satz 8, wenn q irgend ein unzerlegbares, von 0 und 1 verschiedenes Element von $\mathfrak{R}$ ist, in $\mathfrak{B}_\mathfrak{n}$ ein zu $q\mathfrak{n}$ teilerfremdes Ideal $\mathfrak{c}$, für welches $\mathfrak{a}\mathfrak{c}$ ein Hauptideal ist:

$$\mathfrak{a}\mathfrak{c} = \beta\mathfrak{n}, \quad \beta \in \mathfrak{n}.$$

Nach Satz 27 ist $N(\mathfrak{c})$ zu $q\mathfrak{R}$ teilerfremd. Das Primideal $q\mathfrak{R}$ geht also in den Idealen $N(\mathfrak{a})$ und

$$N(\beta\mathfrak{n}) = N(\mathfrak{a})N(\mathfrak{c})$$

zur selben Potenz auf. Da aber β in $\mathfrak{a}$ liegt und von Null verschieden ist, so kann auch im größten gemeinsamen Teiler aller $N(\alpha\mathfrak{n})$ $(\alpha \in \mathfrak{a}, \alpha \neq 0)$ kein von Null- und Einheitsideal verschiedenes Primideal zu höherer Potenz aufgehen als in $N(\mathfrak{a})$. Somit ist $N(\mathfrak{a})$ selbst der größte gemeinsame Teiler.

Satz 29. *Ist* $\mathfrak{n}$ *eine Ordnung mit dem Quotientenkörper* T, $\mathfrak{a}$ *ein vom Nullmodul verschiedener, in* T *gelegener Modul von der Ordnung* $\mathfrak{n}$, *so hat* $\mathfrak{a}^{(i)}$ $(i = 1, \ldots, n)$ *die Ordnung* $\mathfrak{n}^{(i)}$. *Ist* $\mathfrak{a}$ *insbesondere umkehrbar, so ist jedes* $\mathfrak{a}^{(i)}$ *umkehrbar.*

Beweis. $\mathfrak{a}^{(i)}$ liegt im Körper $\mathsf{T}^{(i)}$. Aus

$$\alpha\mathfrak{a}^{(i)} \leqq \mathfrak{a}^{(i)} \qquad (\alpha \in \mathsf{T}^{(i)})$$

folgt, wenn etwa

$$\alpha = \beta^{(i)} \qquad (\beta \in \mathsf{T})$$

gesetzt wird,

$$(\beta\mathfrak{a})^{(i)} \leqq \mathfrak{a}^{(i)},$$
$$\beta\mathfrak{a} \leqq \mathfrak{a},$$
$$\beta \in \mathfrak{n},$$
$$\alpha \in \mathfrak{n}^{(i)}.$$

Umgekehrt folgt aus

$$\alpha \in \mathfrak{n}^{(i)}$$

wieder

$$\alpha\mathfrak{a}^{(i)} \leqq \mathfrak{a}^{(i)}.$$

Der Modul $\mathfrak{a}^{(i)}$ hat also die Ordnung $\mathfrak{n}^{(i)}$.

Ist $\mathfrak{a}$ umkehrbar, also

$$\mathfrak{a}\mathfrak{a}^{-1} = \mathfrak{n},$$

so folgt sofort

$$\mathfrak{a}^{-1} \leqq \mathsf{T},$$
$$\mathfrak{a}^{(i)}(\mathfrak{a}^{-1})^{(i)} = \mathfrak{n}^{(i)},$$

also nach Satz 1 die Umkehrbarkeit von $\mathfrak{a}^{(i)}$.

Satz 30. *Es sei* $\mathfrak{n}$ *eine endliche Ordnung vom Quotientenkörper* T, *und* $\mathfrak{a}$ *sei ein Ideal aus* $\mathfrak{B}_\mathfrak{n}$. *Dann ist*

$$\mathfrak{a}^{(1)} \cdots \mathfrak{a}^{(n)} = N(\mathfrak{a})\mathfrak{n}^{(1)} \cdots \mathfrak{n}^{(n)}.$$

Beweis. Ist α ein beliebiges von Null verschiedenes Element von $\mathfrak{a}$, so ist

$$\mathfrak{a} \supseteq \alpha\,\mathfrak{n},$$

$$\mathfrak{a}^{(1)} \cdots \mathfrak{a}^{(n)} \supseteq (\alpha^{(1)}\,\mathfrak{n}^{(1)}) \cdots (\alpha^{(n)}\,\mathfrak{n}^{(n)}) = \alpha^{(1)} \cdots \alpha^{(n)} \cdot \mathfrak{n}^{(1)} \cdots \mathfrak{n}^{(n)}.$$

Wird $\mathfrak{n}^{(1)} \cdots \mathfrak{n}^{(n)}$ kurz mit $\mathfrak{N}$ bezeichnet, so ist also nach Satz 26

$$\mathfrak{a}^{(1)} \cdots \mathfrak{a}^{(n)} \supseteq N(\alpha\,\mathfrak{n})\,\mathfrak{N}.$$

In $\mathfrak{N}$ ist nach Satz 28 der größte gemeinsame Teiler der $N(\alpha\,\mathfrak{n})$ $(\alpha \in \mathfrak{a},\ \alpha \neq 0)$ durch $N(\mathfrak{a})$ gegeben. Da in $\mathfrak{N}$ der Teilerkettensatz gilt, so stellt man sich mühelos endlich viele von Null verschiedene Elemente $\alpha_1, \ldots, \alpha_r$ von $\mathfrak{a}$ her, für die $N(\mathfrak{a})$ der größte gemeinsame Teiler der Ideale $N(\alpha_1\,\mathfrak{n}), \ldots, N(\alpha_r\,\mathfrak{n})$ ist. Es ist nun zunächst

$$\mathfrak{a}^{(1)} \cdots \mathfrak{a}^{(n)} \supseteq N(\alpha_k\,\mathfrak{n})\,\mathfrak{N} \qquad (k = 1, \ldots, r),$$

also

$$\mathfrak{a}^{(1)} \cdots \mathfrak{a}^{(n)} \supseteq N(\alpha_1\,\mathfrak{n})\,\mathfrak{N} + \cdots + N(\alpha_r\,\mathfrak{n})\,\mathfrak{N} = (N(\alpha_1\,\mathfrak{n}) + \cdots + N(\alpha_r\,\mathfrak{n}))\,\mathfrak{N},$$

$$\mathfrak{a}^{(1)} \cdots \mathfrak{a}^{(n)} \supseteq N(\mathfrak{a})\,\mathfrak{N}. \tag{11}$$

Nach Satz 29 und Satz 3 ist $\mathfrak{a}^{(1)} \cdots \mathfrak{a}^{(n)}$ ein umkehrbarer Modul von der Ordnung $\mathfrak{n}^{(1)} \cdots \mathfrak{n}^{(n)} = \mathfrak{N}$, zugleich offenbar Ideal in $\mathfrak{N}$. Da $N(\mathfrak{a})\,\mathfrak{N}$ ein vom Nullideal verschiedenes Hauptideal in $\mathfrak{N}$ ist, so folgt aus Satz 5 und (11) das Vorhandensein eines Ideals $\mathfrak{X}$ in $\mathfrak{B}_{\mathfrak{N}}$, für welches

$$N(\mathfrak{a})\,\mathfrak{N} = \mathfrak{a}^{(1)} \cdots \mathfrak{a}^{(n)}\,\mathfrak{X} \tag{12}$$

ist. Satz 6 zufolge liegt in $\mathfrak{B}_{\mathfrak{n}}$ ein Ideal $\mathfrak{b}$, für welches $\mathfrak{a}\,\mathfrak{b}$ ein Hauptideal $\mu\,\mathfrak{n}$ $(\mu \in \mathfrak{n})$ ist. Dann ergibt sich in derselben Weise, daß in $\mathfrak{B}_{\mathfrak{N}}$ ein Ideal $\mathfrak{Y}$ mit

$$N(\mathfrak{b})\,\mathfrak{N} = \mathfrak{b}^{(1)} \cdots \mathfrak{b}^{(n)}\,\mathfrak{Y} \tag{13}$$

liegt. Multiplikation von (12) und (13) liefert

$$N(\mathfrak{a})\,N(\mathfrak{b})\,\mathfrak{N} = (\mathfrak{a}\,\mathfrak{b})^{(1)} \cdots (\mathfrak{a}\,\mathfrak{b})^{(n)}\,\mathfrak{X}\,\mathfrak{Y},$$

$$N(\mu\,\mathfrak{n})\,\mathfrak{N} = \mu^{(1)} \cdots \mu^{(n)}\,\mathfrak{N} \cdot \mathfrak{X}\,\mathfrak{Y} = \mu^{(1)} \cdots \mu^{(n)}\,\mathfrak{X}\,\mathfrak{Y},$$

also nach Satz 26

$$\mu^{(1)} \cdots \mu^{(n)}\,\mathfrak{N} = \mu^{(1)} \cdots \mu^{(n)}\,\mathfrak{X}\,\mathfrak{Y},$$

$$\mathfrak{X}\,\mathfrak{Y} = \mathfrak{N},$$

$$\mathfrak{X} = \mathfrak{N},$$

$$N(\mathfrak{a})\,\mathfrak{N} = \mathfrak{a}^{(1)} \cdots \mathfrak{a}^{(n)}\,\mathfrak{N} = \mathfrak{a}^{(1)} \cdots \mathfrak{a}^{(n)}.$$

Satz 31. *Es seien* $\mathfrak{n}_1$ *und* $\mathfrak{n}_2$ *endliche Ordnungen vom Quotientenkörper* T, *und* $\mathfrak{a}$ *und* $\mathfrak{b}$ *seien Ideale aus* $\mathfrak{B}_{\mathfrak{n}_1}$ *bzw.* $\mathfrak{B}_{\mathfrak{n}_2}$. *Dann ist*

$$N(\mathfrak{a})\,N(\mathfrak{b}) = N(\mathfrak{a}\,\mathfrak{b}).$$

Beweis. Nach Satz 30 ist

$$\mathfrak{a}^{(1)} \cdots \mathfrak{a}^{(n)} = N(\mathfrak{a})\, \mathfrak{n}_1^{(1)} \cdots \mathfrak{n}_1^{(n)},$$

$$\mathfrak{b}^{(1)} \cdots \mathfrak{b}^{(n)} = N(\mathfrak{b})\, \mathfrak{n}_2^{(1)} \cdots \mathfrak{n}_2^{(n)},$$

also, wenn $(\mathfrak{n}_1 \mathfrak{n}_2)^{(1)} \cdots (\mathfrak{n}_1 \mathfrak{n}_2)^{(n)}$ kurz mit $\mathfrak{N}$ bezeichnet wird,

$$(\mathfrak{a}\mathfrak{b})^{(1)} \cdots (\mathfrak{a}\mathfrak{b})^{(n)} = N(\mathfrak{a})\, N(\mathfrak{b})\, \mathfrak{N}.$$

Nach Satz 3 gehört $\mathfrak{a}\mathfrak{b}$ dem Bereiche $\mathfrak{B}_{\mathfrak{n}_1 \mathfrak{n}_2}$ an, und $\mathfrak{n}_1 \mathfrak{n}_2$ ist offenbar ebenfalls eine endliche Ordnung vom Quotientenkörper T. Satz 30 liefert also

$$(\mathfrak{a}\mathfrak{b})^{(1)} \cdots (\mathfrak{a}\mathfrak{b})^{(n)} = N(\mathfrak{a}\mathfrak{b})\, \mathfrak{N},$$

$$N(\mathfrak{a})\, N(\mathfrak{b})\, \mathfrak{N} = N(\mathfrak{a}\mathfrak{b})\, \mathfrak{N}.$$

In $\mathfrak{N}$ gibt es Elemente r_1 und r_2 mit $N(\mathfrak{a})\, N(\mathfrak{b}) = r_1 \mathfrak{N}$, $N(\mathfrak{a}\mathfrak{b}) = r_2 \mathfrak{N}$. Es folgt

$$r_1 \mathfrak{N} = r_2 \mathfrak{N},$$

also nach Satz 25

$$N(\mathfrak{a})\, N(\mathfrak{b}) = r_1 \mathfrak{R} = r_1 \mathfrak{N} \cap \mathfrak{R} = r_2 \mathfrak{N} \cap \mathfrak{R} = r_2 \mathfrak{R} = N(\mathfrak{a}\mathfrak{b}).$$

§ 7.

Faktorzerlegung umkehrbarer Ideale.

In diesem Paragraphen ist in $\mathfrak{R}$ der Teilerkettensatz vorausgesetzt: außerdem sei $\mathfrak{R}$ in seinem Quotientenkörper ganz-abgeschlossen[43]). $\mathfrak{n}$ bedeutet durchweg eine endliche Ordnung. Auch in $\mathfrak{n}$ gilt dann der Teilerkettensatz[44]). In $\mathfrak{n}$ möge ferner der Vielfachenkettensatz mod den vom Nullideal verschiedenen Idealen gelten. Endlich sei der Quotientenkörper von $\mathfrak{n}$, der offenbar endlich in bezug auf den Quotientenkörper von $\mathfrak{R}$ ist, von erster Art über diesem.

Satz 32[45]). *Liegt das Primideal* $\mathfrak{p}$ *in* $\mathfrak{B}_\mathfrak{n}$, *so ist in* $\mathfrak{n}$ *jedes zu* $\mathfrak{p}$ *gehörige Primärideal Potenz von* $\mathfrak{p}$.

Beweis. Im Falle $\mathfrak{p} = \mathfrak{n}$ ist die Behauptung klar. Es sei $\mathfrak{p} \neq \mathfrak{n}$. Für jede natürliche Zahl r ist $\mathfrak{p}^{r+1} \subset \mathfrak{p}^r$; denn da $\mathfrak{p}$ in $\mathfrak{B}_\mathfrak{n}$ liegt, so würde aus $\mathfrak{p}^{r+1} = \mathfrak{p}^r = \mathfrak{p}^r \mathfrak{n}$ nach Satz 5, 2. folgen $\mathfrak{p} = \mathfrak{n}$[46]). Ist also $\mathfrak{q}$ ein zu $\mathfrak{p}$

---

[43]) Vermöge der Gültigkeit des Teilerkettensatzes in $\mathfrak{R}$ erstreckt sich die Zwiespältigkeit des Ganzheitsbegriffes (vgl. Fußnote [38])) nicht auf die Ganzheit in bezug auf $\mathfrak{R}$. Vgl. van der Waerden, a. a. O., § 98.

[44]) Nach N., § 2.

[45]) Bei diesem Satz wird lediglich benutzt, daß in $\mathfrak{n}$ der Teilerkettensatz und der Vielfachenkettensatz mod den vom Nullideal verschiedenen Idealen gelten; die obigen schärferen Voraussetzungen über $\mathfrak{R}$ und $\mathfrak{n}$ sind unwesentlich.

[46]) Statt der Zugehörigkeit von $\mathfrak{p}$ zu $\mathfrak{B}_\mathfrak{n}$ kann man an dieser Stelle nach einem Hilfssatz aus N. (§ 8, 1.) auch die Gültigkeit des Teilerkettensatzes in $\mathfrak{R}$ (nebst $\mathfrak{p} \neq (0)$) benutzen.

gehöriges Primärideal, so gibt es, da eine Potenz von $\mathfrak{p}$ durch $\mathfrak{q}$ teilbar ist, eine natürliche Zahl s so, daß $\mathfrak{q}$ durch $\mathfrak{p}^s$, aber nicht durch $\mathfrak{p}^{s+1}$ teilbar ist. Nach Satz 5, 1. gibt es ein Ideal $\mathfrak{a}$ in $\mathfrak{n}$ mit

$$\mathfrak{q} = \mathfrak{p}^s \mathfrak{a}.$$

Wegen $\mathfrak{p} \neq (0)$ ist $\mathfrak{q} \neq (0)$. Somit ist $\mathfrak{a}$ durch kein vom Einheitsideal verschiedenes Primideal teilbar; denn für dieses käme nur $\mathfrak{p}$ in Frage, und aus $\mathfrak{a} \subseteq \mathfrak{p}$ würde folgen $\mathfrak{q} \subseteq \mathfrak{p}^{s+1}$, gegen die Annahme. Also ist $\mathfrak{a} = \mathfrak{n}$, $\mathfrak{q} = \mathfrak{p}^s$.[47])

Satz 33. *Ist* $\mathfrak{o}$ *der Ring der in bezug auf* $\mathfrak{R}$ *ganzen Elemente des Quotientenkörpers von* $\mathfrak{n}$ *und bedeutet* $\mathfrak{f}$ *den Führer von* $\mathfrak{n}$ *nach* $\mathfrak{o}$, *so ist jedes in* $\mathfrak{B}_\mathfrak{n}$ *gelegene Primideal zu* $\mathfrak{f}$ *teilerfremd.*

Beweis. Die Ordnung $\mathfrak{o}$ ist endlich[48]), hat mit $\mathfrak{n}$ denselben Quotientenkörper und ist in diesem ganz-abgeschlossen[49]). Aus einem Satz von Grell[50]) ergibt sich infolgedessen, daß jedes Primideal in $\mathfrak{n}$, dessen zugehörige Primärideale aus ihm sämtlich durch Potenzieren gewonnen werden können, entweder das Einheitsideal ist oder in $\mathfrak{f}$ nicht aufgeht, in beiden Fällen also, falls es sich nicht um das Nullideal handelt, zu $\mathfrak{f}$ teilerfremd ist Zufolge des Satzes 32 ist also jedes in $\mathfrak{B}_\mathfrak{n}$ gelegene Primideal zu $\mathfrak{f}$ teilerfremd.

Satz 34. *Gibt es einen Bereich* $\mathfrak{P}_\mathfrak{n}$ *aus Idealen in* $\mathfrak{n}$ *von der Beschaffenheit, daß alle Ideale aus* $\mathfrak{B}_\mathfrak{n}$ *sich* (*bis auf die Reihenfolge der Faktoren*) *eindeutig als Produkte*[20]) *von Idealen aus* $\mathfrak{P}_\mathfrak{n}$ *darstellen lassen, so ist die Ordnung* $\mathfrak{n}$ *in ihrem Quotientenkörper ganz-abgeschlossen.*

Beweis. Alle bei den Zerlegungen der Ideale von $\mathfrak{B}_\mathfrak{n}$ wirklich vorkommenden Ideale aus $\mathfrak{P}_\mathfrak{n}$ gehören nach Satz 4 zu $\mathfrak{B}_\mathfrak{n}$. Es sei nun $\mathfrak{p}$ ein Ideal aus dem Durchschnitt der Bereiche $\mathfrak{B}_\mathfrak{n}$ und $\mathfrak{P}_\mathfrak{n}$; α und β seien von Null verschiedene Elemente von $\mathfrak{n}$, und $\alpha\beta$ liege in $\mathfrak{p}$. Nach Satz 5 gibt es ein Ideal $\mathfrak{r}$ aus $\mathfrak{B}_\mathfrak{n}$ mit

$$\alpha\mathfrak{n} \cdot \beta\mathfrak{n} = \mathfrak{p}\mathfrak{r}.$$

---

[47]) Nach derselben Methode findet man für alle Ordnungen $\mathfrak{n}$, in denen die beiden Kettensätze gelten, leicht den folgenden Satz: Ist $\mathfrak{q}$ Primärideal aus $\mathfrak{B}_\mathfrak{n}$, $\mathfrak{p}$ das zugehörige Primideal, $\mathfrak{q}_1$ Primärideal in $\mathfrak{n}$ und echtes Vielfaches von $\mathfrak{q}$, so ist $\mathfrak{q}_1$ Vielfaches von $\mathfrak{p}\mathfrak{q}$. Im besonderen Falle $\mathfrak{q} = \mathfrak{p}$ sagt er aus, daß, wenn ein Primideal $\mathfrak{p}$ in $\mathfrak{B}_\mathfrak{n}$ enthalten ist, zwischen $\mathfrak{p}$ und $\mathfrak{p}^2$ kein weiteres Ideal liegt, eine Eigenschaft, aus welcher (nach N., § 8) umgekehrt die nur scheinbar schärfere Aussage des Satzes 32 folgt.

[48]) Vgl. van der Waerden, a. a. O., § 99.

[49]) Vgl. van der Waerden, a. a. O., § 98. Wegen der Endlichkeit von $\mathfrak{o}$ gilt nämlich in $\mathfrak{o}$ der Teilerkettensatz.

[50]) G., S. 553. Der dortige Beweis bleibt wörtlich erhalten, wenn der Vielfachenkettensatz mod den vom Nullideal verschiedenen Idealen nicht, wie dort, in $\mathfrak{R}$ (von wo man nach G., § 1, Satz 1 auf $\mathfrak{n}$ schließen kann), sondern nur, wie oben, in $\mathfrak{n}$ vorausgesetzt wird.

Wegen der vorausgesetzten eindeutigen Faktorzerlegung muß also $\mathfrak{p}$ in $\alpha\mathfrak{n}$ oder $\beta\mathfrak{n}$ aufgehen. Mithin ist jedes zugleich zu $\mathfrak{B}_\mathfrak{n}$ und $\mathfrak{P}_\mathfrak{n}$ gehörige Ideal ein Primideal, also nach Satz 33 teilerfremd zum Führer $\mathfrak{k}$ von $\mathfrak{n}$ nach dem Ring $\mathfrak{o}$ der in bezug auf $\mathfrak{R}$ ganzen Elemente des Quotientenkörpers von $\mathfrak{n}$. Alle Ideale aus $\mathfrak{B}_\mathfrak{n}$ sind Produkte von Idealen aus dem Durchschnitt von $\mathfrak{B}_\mathfrak{n}$ und $\mathfrak{P}_\mathfrak{n}$, also ebenfalls zu $\mathfrak{k}$ teilerfremd. Wegen der Gleichheit der Quotientenkörper von $\mathfrak{n}$ und $\mathfrak{o}$ ist $\mathfrak{k}$ vom Nullideal verschieden[51]). Man greife aus $\mathfrak{k}$ ein von Null verschiedenes Element γ heraus; dann ist das Ideal $\gamma\mathfrak{n}$ Element von $\mathfrak{B}_\mathfrak{n}$, also zu $\mathfrak{k}$ teilerfremd, somit

$$\mathfrak{n} = \gamma\mathfrak{n} + \mathfrak{k} = \mathfrak{k} = \mathfrak{k}\mathfrak{o} = \mathfrak{n}\mathfrak{o} = \mathfrak{o}.$$

[51]) Nach dem Satz aus G., S. 548 f. Vgl. Fußnote [26]).

(Eingegangen am 26. Februar 1931.)

Über wechselseitig affinparallele Flächen.

Von

Wilhelm Süß in Greifswald.

In Analogie zu den gewöhnlichen Parallelflächen nennen wir eine Fläche $\mathfrak{x}$ *affinparallel* zu einer zweiten $\mathfrak{x}^*$, wenn die Punkte von $\mathfrak{x}$ und $\mathfrak{x}^*$ einander derart durch Parallelismus der Tangentenebenen eineindeutig zugeordnet sind, daß die Affinentfernungen der Punkte von $\mathfrak{x}$ von den ihnen entsprechenden von $\mathfrak{x}^*$ konstant sind. Durch die Forderung, daß auch umgekehrt dann $\mathfrak{x}^*$ zu $\mathfrak{x}$ affinparallel sein soll, wird eine gewisse Klasse von *Weingartenflächen der affinen Flächentheorie* bestimmt, die Herr W. van der Woude[1]) und ich[2]) angegeben haben. Ich möchte hier (I.) für den genannten Sachverhalt einen einfachen, die früheren Rechnungen fast ganz umgehenden, auch für Hyperflächen gültigen Beweis angeben, der zugleich den Nutzen der relativen Differentialgeometrie deutlich macht. Sodann (II.) führt die Umkehrung der Betrachtung zu einer anderen Kennzeichnung der betrachteten Flächenpaare.

I. Die Voraussetzungen für die gewöhnlichen Normalenvektoren ξ und die Affinentfernungen E lauten[3])

$$\xi^*(u^1, u^2) = \pm\xi(u^1, u^2, \tag{1}$$

$$\begin{cases} E(\mathfrak{x}, \mathfrak{x}^*) = (\mathfrak{x}^* - \mathfrak{x})\,\mathfrak{X} = p = \text{konst.}, \\ E(\mathfrak{x}^*, \mathfrak{x}) = (\mathfrak{x} - \mathfrak{x}^*)\,\mathfrak{X}^* = p^* = \text{konst.} \end{cases} \tag{2}$$

Nach den Ableitungsgleichungen der Affingeometrie führt der wegen (1), (2) gültige Ansatz

$$\mathfrak{x}^* - \mathfrak{x} = a^i \mathfrak{x}_i + p\,\mathfrak{y}$$

---

[1]) Ein Problem der Affingeometrie. Math. Zeitschr. 26, S. 186–195 (1927).

[2]) Wechselseitig affinparallele Kurven und Flächen. Math. Annalen 98 (1927), S. 313–320

[3]) Wir übernehmen die Bezeichnungen aus W. Blaschke, Differentialgeometrie II. Berlin 1923, besonders § 63.

zu der Relation

$$\mathfrak{x}^* - \mathfrak{x} = p\,\mathfrak{y} = -p^*\mathfrak{y}^*, \tag{3}$$

die wegen $\mathfrak{y}\mathfrak{X} = \mathfrak{y}^*\mathfrak{X}^* = 1$ wieder

$$p\,\mathfrak{X}^* + p^*\mathfrak{X} = 0 \tag{4}$$

zur Folge hat[4]). Wählen wir etwa die sogenannte Relativkrümmungslinien von $\mathfrak{x}^*$ bezüglich $\mathfrak{x}$ als Eichfläche zu Parameterkurven, d. h. dasjenige Netz konjugierter Kurven auf $\mathfrak{x}^*$, dem ein Netz konjugierter Kurven auf $\mathfrak{x}$ entspricht, so wird

$$\mathfrak{x}_i^* = r_i\,\mathfrak{x}_i \quad (i = 1, 2); \qquad G_{12} = G_{12}^* = 0. \tag{5}$$

Hieraus folgt zunächst bei Differentiation von (3) nach den Ableitungsgleichungen

$$1 - \frac{p}{R_i} = r_i \quad (i = 1, 2); \qquad B_{12} = B_{12}^* = 0; \tag{6}$$

die Parameterkurven sind also Affinkrümmungslinien.

Bemerken wir ferner, daß die Längen der Vektoren $\mathfrak{X}$ und $\mathfrak{X}^*$ durch die Gaußschen Krümmungen K bzw. K^* wegen

$$\mathfrak{X} = |K|^{-\frac{1}{4}}\xi, \qquad \mathfrak{X}^* = |K^*|^{-\frac{1}{4}}\xi^* \tag{7}$$

bestimmt sind, daß andererseits wegen (5)

$$r_1 r_2 = \frac{K}{K^*} \tag{8}$$

ist, so lehrt (4) und (6)

$$\left(1 - \frac{p}{R_1}\right)\left(1 - \frac{p}{R_2}\right) = r_1 r_2 = \left(\frac{p^*}{p}\right)^4 = \text{konst.} \tag{9}$$

Das Bestehen der Gleichung (9) *ist eine notwendige Bedingung dafür, daß die Flächen* $\mathfrak{x}$ *und* $\mathfrak{x}^*$ *wechselseitig affinparallel sind. Die Bedingung ist auch hinreichend*, wie die Betrachtungen in II. nebenbei zeigen[1]) [2]).

II. Als notwendige Bedingung ergibt sich aus (9) auch, daß die Relativkrümmung $k = \frac{1}{r_1 r_2}$ von $\mathfrak{x}^*$ bezüglich $\mathfrak{x}$ als Eichfläche konstant sein muß. Wir beweisen nun:

Eine Fläche $\mathfrak{x}$ *und eine zu ihr affinparallele Fläche* $\mathfrak{x}^*$ *sind dann und nur dann wechselseitig affinparallel, wenn die Relativkrümmung der einen bezüglich der anderen als Eichfläche konstant ist.*

Zunächst folgt aus (1), (5) und den beiden ersten Gleichungen (2) und (3) die Gültigkeit von (6), und aus der vorausgesetzten Konstanz von k folgt (9). Durch (9) werde gleichzeitig $|p^*|$ definiert. Machen wir

[4]) Ist (1) und (3) Definition, so folgt daraus wieder (2).

nun den Ansatz

$$p\,\mathfrak{X}^* + q\,\mathfrak{X} = 0,$$

so ist nach (7), (8) und (9): $q = p^* = \text{konst.}$, indem dabei auch das Vorzeichen von p^* bestimmt wird. Da nunmehr (4) gilt, so schließt man aus der ersten Gleichung (3) auf das Bestehen der zweiten Gleichungen (2) und (3). Die anfangs als notwendig erkannte Bedingung hat sich somit auch als hinreichend erwiesen, w. z. b. w.

Für ebene Kurven erhält man analog zu (9) $1 - \frac{p}{R} = \text{konst.}$, die Affinkrümmung $\frac{1}{R}$ muß also konstant sein, was bekanntlich nur für Kegelschnitte der Fall ist. Bei n-dimensionalen Hyperflächen im $(n+1)$-dimensionalen Raum läuft der Index i von 1 bis n; die Exponenten von K und K^* in (7) sind dann $-\frac{1}{n+2}$ und an die Stelle von (9) tritt

$$(9') \qquad \prod_{i=1}^{n}\left(1 - \frac{p}{R_i}\right) = \prod_{i=1}^{n} r_i = \left(\frac{p^*}{p}\right)^{n+2} = \text{konst.}$$

(Eingegangen am 4. März 1931.)

On avoidable Points of Continua with an Application to end Points[1].

By

W. L. Ayres in Ann Arbor, Mich. (USA).

1. In a paper[2] published after his death, Paul Urysohn defines an avoidable point (vermeidbarer Punkt) of a continuum. He calls a point x of a continuum M an avoidable point if for each $\varepsilon > 0$ there exists a $\delta > 0$ so that if y and z are points of $M - x$ such that $\varrho(x, y) < \delta$, $\varrho(x, z) < \delta$, then there exists a continuum $M_{yz} \subset M - x$, containing $x + y$, and lying within a distance ε of x. We shall call such a point a *strong avoidable point* of M and shall call a point a *weak avoidable point* if M_{yz} exists as above except that it be merely a connected set, not a continuum. Somewhat later but before the paper of Urysohn was published, R. G. Lubben[3] gave this same definition for a set M being connected near the point x. Urysohn does not go very far beyond defining the property. He proves (1) that M is connected im kleinen[4] at every strong avoidable point and (2) that a point of a locally connected continuum[4] M is avoidable if and only if it can be ε-separated by a subcontinuum of M. In this paper we shall give some additional properties following from this definition, and in the second part shall show how this property of being avoidable

[1] Communicated to the American Mathematical Society, September 6, 1928. This paper was completed while the author was holding a National Research Fellowship in Mathematics.

[2] Über im kleinen zusammenhängende Kontinua, Math. Annalen **98** (1927), pp. 296—308.

[3] The separation of plane point sets by curves, Bull. Amer. Math. Soc., **32** (1926), p. 200 (abstract).

[4] A set of points M is said to be *connected im kleinen*, or *locally connected* at a point p of M if for each neighborhood $U(p)$ there exists a neighborhood $V(p)$ and a connected subset K of M such that

$$V(p) \cdot M \subset K \subset U(p).$$

may be used to give a new definition of an end point of a continuum. In the first part of the paper we give a number of necessary and sufficient conditions for weak and strong avoidable points. This is followed with an analysis of a continuum in which every point is avoidable (in either sense), and the section is closed with a discussion of the relations between our two avoidable properties and a property defined by Zarankiewicz which is very similar.

Unless otherwise specified, by the continuum M is meant a connected, compact, metric space. For some theorems the continuum M is assumed to lie in a euclidean plane. In these cases by the continuum M is meant a connected, closed point set, and compactness (or boundedness) is not to be assumed unless stated.

I. On avoidable points.

2. Theorem 1. *In order that a point p of a continuum M be an avoidable point of M in the weak sense it is necessary and sufficient that* (a) *M be locally connected at p*, (b) *p not be a cut point im kleinen*[5]) *of M.*

The conditions are necessary. The necessity of condition (a) has been demonstrated by Urysohn[6]). Now suppose there exists an $\varepsilon > 0$ such that p is a cut point of the component K of $M \cdot S(p, \varepsilon) \supset p$. Then

$$K - p = K_1 + K_2, \qquad K_1 \overline{K}_2 + \overline{K}_1 K_2 = 0.$$

No connected subset $M - p$ which lies in $S(p, \varepsilon)$ contains a point of both K_1 and K_2. Hence there exists no number $\delta > 0$ such that every two points of $(M - p) \cdot S(p, \delta)$ lie together in a connected subset of $(M - p) \cdot S(p, \varepsilon)$. But this is contrary to our assumption that p is a weak avoidable point of M. Thus p is not a cut point im kleinen of M.

The conditions are sufficient. By condition (a) for each neighborhood $U(p)$ there exists a neighborhood $V(p)$ and a connected subset H of M such that

$$M \cdot V(p) \subset H \subset U(p).$$

From this it follows that if K is the component of $M \cdot U(p) \supset p$, then

$$M \cdot V(p) \subset H \subset K.$$

---

[5]) A point p of a connected set M is said to be a *cut point im kleinen* of M if there exists a neighborhood $U(p)$ such that p is a cut point of the component of $M \cdot U(p) \supset p$.

[6]) P. Urysohn, loc. cit. Urysohn proves that if p is an avoidable point of M in the *strong sense*, then M is locally connected at p; but a very slight modification of his argument gives the result we desire.

From condition (b) the set $K - p$ is connected. Hence for each neighborhood $U(p)$ there exists a neighborhood $V(p)$ and a connected set $K - p$ such that

$$M \cdot V(p) - p \subset K - p \subset (M - p) \cdot U(p).$$

Therefore p is a weak avoidable point of M.

3. Since a connected open subset of a locally connected continuum is arc-wise connected, if p is a weak avoidable point of a locally connected continuum it is a strong avoidable point. Thus for locally connected continua the two properties are equivalent and we have the following

Corollary 1 A. *In order that a point p of a locally connected continuum M be an avoidable point of M in the weak sense or in the strong sense it is necessary and sufficient that p not be a cut point im kleinen of M.*

4. Theorem 2. *In order that a point p of a continuum M be a strong avoidable point of M it is necessary and sufficient that* (a) *M be locally connected at p,* (b) *p not be a weak cut point im kleinen*[7]) *of M.*

Condition (a) is necessary as before. Let $U(p)$ be a neighborhood of p and let K denote the component of $M \cdot U(p) \supset p$. Since p is a strong avoidable point of M, there exists a neighborhood $V(p)$ such that any two points of $M \cdot V(p) - p$ lie in a closed and connected subset of $(M - p) \cdot U(p)$. Let x and y be two points of $\overline{K} - p$, and let $W(p)$ be a neighborhood of p such that $\overline{W(p)} \subset V(p)$ and $W(p) \cdot (x + y) = 0$. Let K_x and K_y denote the components of $\overline{K} - W(p)$ containing x and y respectively. Let x' and y' be points of $K_x \cdot [V(p) - W(p)]$ and $K_y \cdot [V(p) - W(p)]$ respectively. Since $x' + y' \subset V(p) - p$, there exists a continuum H such that

$$x' + y' \subset H \subset [M - p] \cdot U(p).$$

As K is the component of $M \cdot U(p) \supset x' + y'$, we have $H \subset K$. Then $K_x + K_y + H$ is a continuum and

$$x + y \subset K_x + K_y + H \subset \overline{K} - p.$$

Hence $\overline{K} - p$ is a semi-continuum and p is not a cut point im kleinen.

We will now show that the condition is sufficient. Let $U(p)$ be any neighborhood of p. By condition (b) there exists a neighborhood $V(p) \subset U(p)$ such that if K is the component of $M \cdot V(p) \supset p$, then $\overline{K} - p$ is a semi-continuum. From condition (a) there exists a neighborhood $W(p)$ such that $M \cdot W(p) \subset K$. Then for each pair of points x and y of $M \cdot W(p) - p$

[7]) A point p of a continuum M is said to be a *weak cut point im kleinen* if there exists a neighborhood $U(p)$ such that if the neighborhood $V(p) \subset U(p)$, and K is the component of $M \cdot V(p) \supset p$, then $\overline{K} - p$ is not a semi-continuum.

there exists a continuum M_{xy} such that

$$x + y \subset M_{xy} \subset \overline{K} - p \subset \overline{V(p)}.$$

Thus p is a strong avoidable point.

5. Theorem 3. *In order that a point p of a plane continuum M be a weak avoidable point of M it is necessary and sufficient that* (a) *M be locally connected at p, and* (b) *if C is any circle with center at p then p be not accessible from more than one complementary domain of $M + C$.*

The first condition is necessary as in Theorem 1. Suppose there exists a circle C with center p such that p is accessible from two complementary domains D_1 and D_2 of $M + C$. There exist arcs α and β which are subsets of $D_1 + p$ and $D_2 + p$ respectively and have p as one end point. Let C_1 be a circle with center p such that $C_1 \cdot \alpha \neq 0 \neq C_1 \cdot \beta$. The arcs α and β contain subarcs xp and yp respectively such that $x + y \subset C_1$ and $[xp - x] + [yp - y] \subset I(C_1)$, the interior of C_1. Then

$$I(C_1) - [xp + yp] = R_1 + R_2, \qquad \overline{R}_1 R_2 + \overline{R}_2 R_1 = 0.$$

Let K denote the component of $M \cdot I(C_1) \supset p$. Since $\langle xp \rangle$ and $\langle yp \rangle$[8]) belong to different complementary domains of $M + C_1$, it follows that both R_1 and R_2 contain points of K. Let $K_1 = R_1 \cdot K$ and $K_2 = R_2 \cdot K$. Then $\overline{K}_1$ and $\overline{K}_2$ are continua such that

$$\overline{K} = \overline{K}_1 + \overline{K}_2 \quad \text{and} \quad \overline{K}_1 \cdot \overline{K}_2 = p.$$

Then p is a cut point of K and is a cut point im kleinen of M. But this is impossible by Theorem 1[9]).

In view of Theorem 1, to prove the conditions sufficient we have only to show that under the assumed conditions p is not a cut point im kleinen of M[9]). Suppose that it is. Then there exists a circle C with center p such that if K denotes the component of $M \cdot [C + I(C)] \supset p$, then

$$K - p = K_1 + K_2, \quad \overline{K}_1 \cdot K_2 + \overline{K}_2 \cdot K_1 = 0.$$

Evidently we may choose C of radius such that there exist points

$$s \in K_1 \cdot C \quad \text{and} \quad t \in K_2 \cdot C.$$

By a theorem of R. G. Lubben[10]) there exists a simple closed curve J separating s and t and such that $J \cdot K = p$. By condition (a) there exists

---

[8]) If abc is an arc with end points a and c, the symbols $\langle abc, abc \rangle$, and $\langle abc \rangle$ denote the sets $abc - a$, $abc - c$, and $abc - a - c$ respectively.

[9]) While Theorem 1 does not apply directly since M is not compact, every continuum in the plane is locally compact and this is sufficient for Theorem 1.

[10]) Separation theorems with applications to questions concerning accessibility and plane continua, Trans. Amer. Math. Soc., **31** (1929), pp. 503—522, theorem 18.

a $\delta > 0$ such that $M \cdot S(p, \delta) \subset K$. Hence J contains two arcs $x\,p$ and $y\,p$ such that

$$(x\,p) \cdot (y\,p) = (x\,p + y\,p) \cdot (M + C) = p.$$

By condition (b) the points x and y must belong to the same complementary domain of $M + C$. There exists an arc $x\,z\,y$ lying entirely in this domain. Now from $x\,z\,y$ and the component of $J \cdot I(C) \supset p$ we may obtain a simple closed curve J_1 separating s and t. Then either s or t belongs to $I(J_1)$. But this is a contradiction for

$$I(J_1) \subset I(C) \quad \text{and} \quad s + t \subset C.$$

Theorem 4. *In order that a point p of a plane continuum M be a weak avoidable point of M it is necessary and sufficient that* (a) *M be locally connected at p,* (b) *p be a non-cut point of M,* (c) *p be accessible from not more than one complementary domain of M.*

Theorem 4 may be derived from Theorem 3 without difficulty.

Corollary 4A. *In order that a point p of a plane locally connected continuum M be a weak avoidable point of M it is necessary and sufficient that p be a non-cut point of M which lies on the boundary of at most one complementary domain of M.*

Corollary 4B. *If the point p of the plane continuum M belongs to the boundary of no complementary domain of M, then p is a weak avoidable point of M if and only if M is locally connected at p.*

6. Theorem 5. *If for each $\varepsilon > 0$ there exists a $\delta > 0$ such that no complementary domain of the plane continuum M of diameter $> \varepsilon$ has a point within a distance δ of p, then p is a strong avoidable point of M.*

Theorem 5 may be proved with a slight extension of a theorem of G. T. Whyburn[11]).

Corollary 5A. *If p is a point of the plane continuum M and* (a) *for each $\varepsilon > 0$ there are only a finite number of complementary domains of M of diameter $> \varepsilon$,* (b) *p belongs to the boundary of no complementary domain of M, then p is a strong avoidable point of M.*

That this result does not remain true without condition (a) even when M is locally connected at p is seen by the following simple example: Let M consist of the intervals from $(1, 1)$ to $(1, -1)$, from $(0, 1)$ to $(0, -1)$, from $(-1, 1)$ to $(-1, -1)$, from $(1, 0)$ to $(-1, 0)$, together with the set of intervals from $(1, 1/n)$ to $(-1, 1/n)$ for all positive and negative integers n. Let p be the point $(0, 0)$. Then M is locally connected at p,

[11]) Concerning the complementary domains of continua, Annals of Math., 29 (1928), pp. 399—411, theorem 1.

and p belongs to the boundary of no complementary domain of M, but p is not a strong avoidable point. However p is a weak avoidable point by Corollary 4 B.

7. Theorem 6. *If every non-cut point of the continuum M is a weak avoidable point, then M is locally connected at every point.*

From Theorem 1 it follows that M is locally connected at each non-cut point. But a continuum which is locally connected at every non-cut point is locally connected at every point.

Theorem 7. *If M lies in the plane and satisfies the hypothesis of Theorem 6, then no two complementary domains of M have more than a countable number of boundary points in common and every such point is a cut point of M.*

Let B_1 and B_2 be the boundaries of two complementary domains of M. Only a countable number of points of $B_1 \cdot B_2$ are cut point of M.[12]) Hence if $B_1 \cdot B_2$ is uncountable, it contains a point which is a non-cut point of M. By hypothesis this point is a weak avoidable point which is impossible by Corollary 4 A.

8. Continua in which every point is avoidable.

If every point of a continuum M is an avoidable point in the weak or in the strong sense, then M is locally connected at every point. And in a locally connected continuum the two types of avoidable points are identical. Hence in this section we may speak merely of avoidable points without distinguishing the senses.

Lemma 8 A. *If every point of K is a weak avoidable point of the continuum M, $M - K$ is the sum of two mutually separated sets M_1 and M_2, and every point of K is a limit point of both M_1 and M_2, then the set K is dense in itself.*

Suppose there exists a point p of K and a number $\varepsilon > 0$ such that $[K - p] \cdot S(p, \varepsilon) = 0$. Let H be the component of $M \cdot S(p, \varepsilon) \supset p$. By Theorem 1 M is connected im kleinen at p and there exists $\delta > 0$ such that $M \cdot S(p, \delta) \subset H$. Let $x \in M_1 \cdot S(p, \delta)$ and $y \in M_2 \cdot S(p, \delta)$. Then $H \supset x + y$. By Theorem 1, $H - p$ is connected. Then $H - p$ is a connected set of $M - K$ containing a point x of M_1 and a point y of M_2, which is impossible as M_1 and M_2 are mutually separated.

Lemma 8 B. *If K is a countable subset of the continuum M and every point of K is a weak avoidable point of M, then $M - K$ is connected.*

---

[12]) G. T. Whyburn, Concerning the cut points of continua, Trans. Amer. Math. Soc. **30** (1928), pp. 597—609, theorems 7 and 10.

Suppose

$$M - K = H + G, \quad \overline{H} \cdot G + \overline{G} \cdot H = 0.$$

Let

$$H_1 = \overline{H} - \overline{G}, \quad G_1 = \overline{G} - \overline{H}, \quad K_1 = \overline{G} \cdot \overline{H}.$$

Then

$$M - K_1 = H_1 + G_1,$$

$$\overline{H}_1 G_1 + G_1 H_1 \subset \overline{H}(\overline{G} - \overline{H}) + \overline{G}(\overline{H} - \overline{G}) = 0.$$

Then as $K_1 \subset K$, the conditions of Lemma 8A are satisfied and K_1 is dense in itself. By definition, K_1 is closed. Hence K_1 is a countable perfect set, which cannot exist.

From this lemma we have immediately

Theorem 8. *If every point of the continuum M is avoidable, then no countable subset of M cuts M.*

Corollary 8A. *If every point of a continuum is avoidable, then every point is of Menger-Urysohn order c.*[13])

This result is of interest in view of the fact that the avoidable points may be uncountable and everywhere dense on the continuum and not one be of order c. An example is an acyclic locally connected continuum in which the end points are everywhere dense. Several such examples have been constructed[14]). Here every avoidable point is an end point and thus of order 1.

In a recent paper[15]) I have shown that if no finite subset of a locally connected continuum M cuts M, then every finite subset may be joined in M by an arc and by a simple closed curve. Then with Theorem 8 we have

Theorem 9. *If every point of the continunm M is avoidable and K is any finite subset of M, there exists a simple closed curve of M containing K (and thus also an arc containing K).*

Recently[16]) I have proved that if the set of all non-cut points of order 2 of a bounded plane locally connected continuum M contains no continuum then every arc of M is extendible to a maximal arc, *i. e.* every

---

[13]) A point of a set K is said to be of Menger-Urysohn order c if it cannot be ε-separated by a subset of K of cardinal number less than c.

[14]) For example, T. Wazewski, Annales Soc. Pol. Math. 2 (1923), pp. 49—170; K. Menger, Fund. Math. 10 (1926), p. 108; W. L. Ayres, Annals of Math. 28 (1927), pp. 405—406. As a result of a typographical error the set M as defined in my paper is not closed. The set $\overline{M}$ is the desired example.

[15]) On joining subsets of a continuous curve by arcs and simple closed curves, not as yet published. For a short abstract see the Anzeiger der Akad. der Wiss. in Wien, Math.-Naturwiss. Klasse, 1929, Nr. 17.

[16]) Conditions under which every arc of a continuous curve is a subset of a maximal arc of the curve, Math. Annalen 101 (1829), pp. 194—209, theorem II.

arc of M is a subset of an arc of M which is a proper subset of no arc of M. Then using Corollary 8 A we have

Theorem 10. *If every point of the bounded plane continuum M is avoidable, then every arc of M is a subset of a maximal arc of M.*

Theorem 11. *In order that every point of a plane continuum M be avoidable it is necessary and sufficient that M be a locally connected continuum which is cyclicly connected and such that no two complementary domains of M have a boundary point in common.*

Suppose every point of M is avoidable. Then the conditions of Theorem 4 are satisfied at every point of M. From condition (a), M is locally connected; from condition (b), M is cyclicly connected[17]). Since M is a plane locally connected continuum, if any two complementary domains have a point in common, then this point is accessible from the two complementary domains. But this is impossible from condition (c) of Theorem 4. Conversely, if our conditions are satisfied it is easy to show that the conditions of Theorem 4 hold at every point and thus every point is avoidable.

Theorem 12. *If every point of the plane bounded continuum M is avoidable and K denotes the sum of the boundaries of all complementary domains of M, then $M - K$ is arc-wise connected and arc-wise connected im kleinen.*

From Theorem 11, M is cyclicly connected and thus the boundary of every complementary domain of M is a simple closed curve[18]). Let G be a collection defined as follows: (1) Let J^* be the boundary of any bounded complementary domain of M. Then J^* plus its interior is an element of G. (2) Every point of M which belongs to the boundary of no complementary domain of M is an element of G. (3) Every element of G is of one of the two preceding types. From Theorem 11 we see that no two elements of G have a point in common. The collection G is upper semi-continuous[19]) and the hyperspace H in which the elements of G are considered as points is equivalent to the euclidean plane. There are only a countable number of elements of G of type (1). Hence $M - K$ corresponds in the hyperspace to the plane minus a countable set. Such a set is arc-wise connected and arc-wise connected im kleinen. But the elements of G corresponding to $M - K$ are points and thus any arc of such elements is an ordinary arc. Therefore $M - K$ has the desired properties.

---

[17]) G. T. Whyburn, Cyclicly connected continuous curves, Proc. Nat. Acad. Sciences (Washington) **13** (1927), pp. 31—38, theorem 1.

[18]) See theorem 10 of paper referred to in footnote [17]).

[19]) R. L. Moore, Concerning upper semi-continuous collections of continua, Trans. Amer. Math. Soc. **27** (1925), pp. 416—428.

Theorem 13. *If every point of a plane bounded continuum M is avoidable, then every point of M may be ε-separated by an arc or a simple closed curve of M according as the point does or does not belong to the boundary of a complementary domain of M.*

If the point p of M belongs to the boundary of no complementary domain of M, it may be ε-separated by a simple closed curve of M.[20]) Now suppose p belongs to the boundary of the complementary domain D of M. The point p may be ε-separated by a single continuum[2]). Since p belongs to the boundary of D, for ε sufficiently small the point p cannot belong to a bounded complementary domain of this ε-separating continuum. Then by a lemma in my paper, "Continuous curves in which every arc may be extended"[21]), the point p may be ε-separated by an arc of M.

9. C. Zarankiewicz[22]) has defined a "point de coupure locale" of a continuum. This is somewhat similar to a point being unavoidable. We may define M to have the *weak* (*strong*) *Zarankiewicz property* at p if for every neighborhood $U(p)$ there exists a neighborhood $V(p) \subset U(p)$ and such that $M \cdot V(p) - p$ is connected (a semi-continuum). Then whenever p is not a "point de coupure locale" of M the continuum M has the strong Zarankiewicz property at p and conversely. These Zarankiewicz properties are very similar to the avoidable properties and it will be interesting to see the relations between the four properties.

Theorem 14. *If the continuum M has the strong Zarankiewicz property at a point p, then it has the weak Zarankiewicz property at p and p is a strong avoidable point of M. If either* (1) *p is a strong avoidable point of M or* (2) *M has the weak Zarankiewicz proverty at p, then p is a weak avoidable point of M. No further relations exist between these four properties.*

Obviously if either strong property exists at p, the corresponding weak property exists at p. We shall show that if M has the weak Zarankiewicz property at p, then p is a weak avoidable point. Let $U(p) = S(p, \varepsilon)$. There exists a neighborhood $V(p) \subset U(p)$ such that $M \cdot V(p) - p$ is connected. There exists a $\delta > 0$ such that $S(p, \delta) \subset V(p)$. Let x and y

---

[20]) G. T. Whyburn, Concerning continua in the plane, Trans. Amer. Math. Soc. 29 (1927), p. 381, theorem 7.

[21]) Bull. Amer. Math. Soc. **35** (1929), p. 852. This lemma is as follows: If the point p of the plane locally connected continuum M may be ε-separated by a set S of n mutually exclusive continua such that p is not contained in a bounded complementary domain of any one, but not by any set S of less than n, then p may be ε-separated by a set of n mutually exclusive arcs.

[22]) Sur les coupures locales faites par les continus, Bull. Acad. Polonaise des Sciences et des Lettres 1927, pp. 193—218.

be points of $[M - p] \cdot S(p, \delta)$. Then

$$x + y \subset M \cdot V(p) - p \subset S(p, \varepsilon).$$

Hence p is a weak avoidable point of M. If M has the strong Zarankiewicz property at p, then $M \cdot V(p) - p$ is a semi-continuum and there exists a continuum M_{xy} such that

$$x + y \subset M_{xy} \subset M \cdot V(p) - p \subset S(p, \varepsilon).$$

Thus if M has the strong Zarankiewicz property at p, the point p is a strong avoidable point of M. We may see that no further relations exist from the following examples:

Example 1. Let M denote the intervals from $(0, 0)$ to $(1, 0)$ and $(0, 1)$ together with the intervals from $(0, 1/n)$ to $(1, 1/n)$ for all posi-

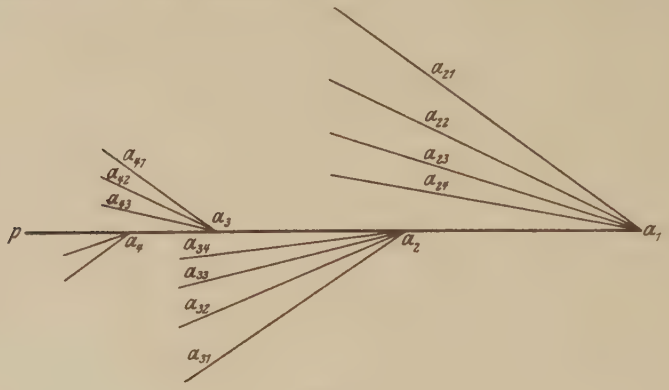


Fig. 1.

tive integers n. Then the point $(0, 0)$ is a weak avoidable point and M has the weak Zarankiewicz property here but neither of the strong properties are true at $(0, 0)$.

Example 2. (See Figure 1.) Let K denote the interval from $(0, 0)$ to $(2, 0)$. Let M_1 denote the set of intervals from $(2, 0)$ to $(1, 1/i)$ for each positive integer i. Let $M_{n+1} = T(M_n)$ $(n = 1, 2, \ldots)$ where T is the transformation

$$x' = 2x/3, \quad y' = -y.$$

Let

$$M = K + \sum_{n=1}^{\infty} M_n.$$

It is easy to see that the point p $(0, 0)$ is a strong avoidable point of M. We will see now that M does not have the weak Zarankiewicz property at p. (p is a „point de coupure locale“ of M.) Let a_1 be the point $(2, 0)$ and $a_n = T(a_{n-1})$ $(n > 1)$. Let a_{2i} be a point of the interval from $(2, 0)$ to $(1, 1/i)$ such that a_2 is the sequential limit point of the

sequence $[a_{2i}]$. Let $a_{ni} = T(a_{n-1,i})$. Now let $U(p)$ be any neighborhood of p that does not contain a_1 and take any neighborhood $V(p) \subset U(p)$. In the order from p to a_1 let q be the first point of $M - V(p)$. Let m be the smallest integer such $a_m \subset pq - q$. As $a_m \subset V(p)$, there is some integer k so that $a_{mk} \subset V(p)$. Every connected subset of M containing a_{mk} and a_m contains the arc $a_{mk}a_{m-1}a_m$. But this arc contains q and thus does not belong to $V(p)$. Then $M \cdot V(p) - p$ is not connected.

II. On end points.

10. We shall apply now the preceding concept of an avoidable point to give a new definition of an end point. We shall consider the two following properties of a point p of a continuum M: (1) the point p is a weak avoidable point of every subcontinuum of M, (2) the point p is a strong avoidable point of every subcontinuum of M. This would seem to give two types of points p but we shall show in Theorem 16 that these two properties are equivalent. This forms a quite interesting result for the two avoidable properties are themselves not equivalent in a general continuum. In the following theorem we shall prove that a point of a locally connected continuum is an end point if and only if it has one of the above two equivalent properties; — thus giving a new definition of an end point. Following this we shall show the relations of our new end point property in an arbitrary continuum to those discussed by H. M. Gehman [23]). Finally we shall treat two problems left unsolved in Gehman's discussion.

11. Theorem 15. *If the point p of the continuum M is a weak avoidable point of every subcontinuum of M, then p is contained in no subcontinuum of M which is irreducible between two other points of M.*

Suppose M contains a subcontinuum $N \supset p$ and irreducible between p' and p'' $(p' \neq p \neq p'')$. For each positive integer i let $r_i = \varrho(p, p' + p'')/2^i$, and let D_i denote the component of $N \cdot S(p, r_i) \supset p$. Let K_i' and K_i'' be the components of $N - D_i$ containing p' and p'' respectively. As p is a weak avoidable point of N, N is locally connected at p; and there exists $\delta_i > 0$ such that $N \cdot S(p, \delta_i) \subset D_i$. Then $[K_i' + K_i''] \cdot p = 0$ for $[K_i' + K_i''] \cdot S(p, \delta_i) = 0$.

If $K_i' \cdot K_i'' \neq 0$, then $\overline{K_i'} + \overline{K_i''}$ is a proper subcontinuum of N containing p' and p''. But by assumption N is an irreducible continuum between p' and p''. Thus $K_i' \cdot K_i'' = 0$ and $K_i' + K_i'' + \overline{D}_i = N$ for every i. Further $\overline{D}_{i+1} \subset \overline{D}_i$ and $K_{i+1}' - K_i' \subset \overline{D}_i$. Thus $K_i' + \overline{D}_i$ is a continuum

[23]) Concerning end points of continuous curves and other continua, Trans. Amer. Math. Soc. **30** (1928), pp. 63—84.

containing p and p' and $K'_{i+1} + \overline{D}_{i+1} \subset K'_i + \overline{D}_i$ for every i. Then

$$K' = \prod_{i=1}^{\infty} (K'_i + \overline{D}_i)$$

is a continuum containing p and p'. Similarly

$$K'' = \prod_{i=1}^{\infty} (K'_i + \overline{D}_i)$$

is a continuum containing p and p''. Now

$$K' \cdot K'' = \prod_{i=1}^{\infty} (K'_i + \overline{D}_i) \cdot \prod_{i=1}^{\infty} (K''_i + \overline{D}_i)$$

$$= \prod_{i=1}^{\infty} [(K'_i + \overline{D}_i)(K''_i + \overline{D}_i)] = \prod_{i=1}^{\infty} \overline{D}_i = p.$$

Then p is not a weak avoidable point of the subcontinuum $K' + K''$ of M for it is a cut point. But it is a weak avoidable point by hypothesis. Thus a contradiction results from the assumption that M contains a subcontinuum containing p and irreducible between two other points.

Theorem 16. *If the point p of the continuum M is a weak avoidable point of every subcontinuum of M, then p is a strong avoidable point of every subcontinuum of M.*

Let N be any subcontinuum of M and ε be any positive number. As p is a weak avoidable point of N there exists a $\delta > 0$ such that if $x + y \subset [N - p] \cdot S(p, \delta)$, there exists a connected set H such that

$$x + y \subset H \subset [N - p] \cdot S\left(p, \tfrac{1}{2}\varepsilon\right).$$

The compact continuum $\overline{H}$ contains a continuum G irreducible between x and y. From the preceding theorem the continuum G does not contain p. Then for any $\varepsilon > 0$ there exists a $\delta > 0$ such that if $x + y \subset [N - p] \cdot S(p, \delta)$, there exists a continuum G such that

$$x + y \subset G \subset [N - p] \cdot \overline{S\left(p, \tfrac{1}{2}\varepsilon\right)} \subset S(p, \varepsilon).$$

Therefore p is a strong avoidable point of N.

12. Theorem 17. *In order that a point p of a locally connected continuum M be an end point it is necessary and sufficient that p be a weak (or strong) avoidable point of every subcontinuum of M.*

In the paper of Gehman[23]) nine equivalent definitions of an end point of a locally connected continuum are listed. As these are all equivalent we may take whichever is most convenient for our purposes here. To prove the necessity of our condition we will take the Menger-Urysohn definition of an end point. Let N be any subcontinuum of M and ε be

any positive number. Since p satisfies the Menger-Urysohn definition there exists a point x of M such that

$$M - x = M_1 + M_2, \quad \overline{M}_1 M_2 + \overline{M}_2 M_1 = 0, \quad M_1 \supset p, \quad M_1 + x \subset S(p, \varepsilon).$$

Choose $\delta > 0$ so that $M_2 \cdot S(p, \delta) = 0$. Now let p' and p'' be points of $[N - p] \cdot S(p, \delta)$. If this last set is vacuous, then p satisfies the definition for a weak (or strong) avoidable point of N vacuously. Now again from the Menger-Urysohn definition there exists a point y such that

$$M_1 + x - y = M_{11} + M_{12}, \quad \overline{M}_{11} M_{12} + \overline{M}_{12} M_{11} = 0, \quad M_{11} \supset p,$$
$$M_{12} \supset p' + p'' + x.$$

It is easy to see that $N \cdot (M_{12} + y)$ is a continuum such that

$$p' + p'' \subset N \cdot (M_{12} + y) \subset S(p, \varepsilon).$$

Then p is a strong avoidable point of N.

In order to see the sufficiency of our condition we take the Whyburn definition of an end point. If p is not an end point, there exists an arc which has p as an interior point. Then p is a cut point of the arc and thus cannot be an avoidable point of it. But by hypothesis p is an avoidable point of every subcontinuum of M.

Corollary 17A. *If every subcontinuum of the locally connected continuum M has the weak (or strong) Zarankiewicz property at p, then p is an end point of M.*

13. In the preceding corollary we have a sufficient condition for an end point which is obtained from the Zarankiewicz properties. We may see that this condition is not necessary in the following example:

For each integer $n > 0$ let C_n be the circle with center $(3/2^n, 0)$ and radius $1/2^n$. Let q be the point $(0, 0)$. Let

$$M = q + \sum_{n=1}^{\infty} (C_n + I(C_n)).$$

Then q is an end point of M since two consecutive circles C_n have just one point in common and q is a point of order 1. We shall now show that M contains a subcontinuum K which has neither the Zarankiewicz weak nor the Zarankiewicz strong property at q. For each positive integer n let x_n be the point $(2^{3-n}, 0)$. Now for each integer $n > 0$ let K_n be a set homeomorphic with the set M of example 2 in section 9 such that a) x_n and x_{n+1} correspond to the points a_1 and p of the example respectively, and (b) $K_n - x_n - x_{n+1} \subset I(C_n)$. Let

$$K = q + \sum_{n=1}^{\infty} K_n.$$

By an argument very similar to that used in example 2 of section 9, we may see that K does not have the Zarankiewicz weak or strong property at q.

14. In a recent article[23]) H. M. Gehman has listed a number of properties used by various authors to define an end point and has studied the relations between these properties in a continuum M. For reference we will list these properties here using the same numbers as in Gehman's paper. We shall list our new definition as property 10.

Property 3. (Modified Wilder.) If N is any subcontinuum of M containing p, the point p is a component of $M - N + p$.

Property 4. (Moore) The point p is not a cut point of any subcontinuum of M.

Property 5. (Yoneyama) The point p is not contained in any subcontinuum of M irreducible between two other points of M.

Property 6. (Modified Wilder) If N is any subcontinuum of M containing p, then p is not a limit point of any component of $M - N$

Property 7. (Menger-Urysohn) The point p is a point of order 1 of M.

Property 8. (Yoneyama) If K_i $(i = 1, 2)$ is a subcontinuum of M irreducible between p and some point p_i, then there exists a continuum H such that

$$p \subset H \subset K_1 \cdot K_2.$$

Property 9. (Yoneyama-Young). There exists an $\varepsilon > 0$ such that if p_i $(i = 1, 2)$ is a point of $M \cdot S(p, \varepsilon)$ and K_i is a subcontinuum of M irreducible between p and p_i, then either $K_1 \subset K_2$ or $K_2 \subset K_1$.

Property 10. (Ayres.) The point p is a weak (or strong) avoidable point of every subcontinuum of M.

Gehman proves the following relations between these properties (see Fig. 2): 7 implies 3, 5 plus 9 implies 3, 3 implies 6, 6 implies 5, 5 implies 4, 8 implies 4, 9 implies 8. He shows by examples that no further relations exist between properties 3 — 9 except possibly that 6 implies 3 or 6 plus 8 implies 3. But in example 7 of this section we may see that neither of these relations is true. Hence all relations between properties 3 — 9 are given by Gehman.

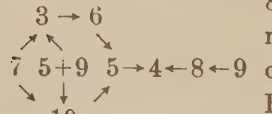


Fig. 2.

We shall now discuss the relations between properties 3 — 10.

Theorem 18. *If a point of a continuum M has property* 7, *it has property* 10.

Theorem 19. *If a point of a continuum M has property* 10, *it has property* 5.

Theorem 18 was proved in showing the necessity of the condition in Theorem 17. In this no use was made of the fact that M was a locally connected continuum. Theorem 19 is exactly the same as Theorem 15.

Theorem 20. *If a point of the continuum M has properties* 5 *and* 9, *it has property* 10.

Let N be any subcontinuum of M containing p and let ε be any positive number. Denote the component of $N \cdot S(p, \varepsilon) \supset p$ by K. Under the hypothesis of our theorem H. M. Gehman[24]) has shown that

$$p \text{ non-}\varepsilon \overline{M - \overline{K}},$$

and thus

$$p \text{ non-}\varepsilon \overline{N - \overline{K}}.$$

Then there exists a positive $\delta < \varepsilon$ such that $[N - \overline{K}] \cdot S(p, \delta) = 0$. Then $N \cdot S(p, \delta) \subset K$. Our theorem is then proved if $K - p$ is connected. If we assume that $K - p$ is not connected, we may easily reach a contradiction by making use of the fact that p has property 5 or that it has property 9.

The following set of examples shows that no further relations that involve property 10 exist between properties 3—10.

Example 1[25]). Let M consist of the point $(0, 0)$ and the circles with center at $(3/2^n, 0)$ and with radius $1/2^n$, for $n = 1, 2, 3, \ldots$, and let p be the point $(0, 0)$. The point p has properties 3—7 and 10, but not 8 and 9.

Example 2[25]). Let M consist of the intervals from $(0, 0)$ to $(1, 0)$, and from $(1/2^n, 0)$ to $(1/2^n, 1/2^n)$ for $n = 1, 2, \ldots$; and let p be the point $(0, 0)$. The point p has properties 3—8 and 10, but not property 9.

Example 3[25]). Let M consist of the curve $y = \sin(1/x)$, for $0 < x \leqq 1$, and the interval from $(0, -1)$ to $(0, 1)$. Let p be the point $(0, 1)$, which has properties 4, 8, and 9, but not 3, 5—7 nor 10.

Example 4[25]). Let M consist of the intervals from $(0, 0)$ to $(1, 0)$, from $(1/2^n, 0)$ to $(1/2^n, -1/2^n)$, from $(1/2^n, -1/2^n)$ to $(-1/2^n, -1/2^n)$, from $(-1/2^n, -1/2^n)$ to $(-1/2^n, 1/2^n)$, from $(-1/2^n, 1/2^n)$ to $(1, 1/2^n)$, from $(1, 1/2^n)$ to $(1, 3/2^{n+2})$, from $(1, 3/2^{n+2})$ to $(0, 3/2^{n+2})$, for $n = 1, 2, \ldots$. Let p be the point $(0, 0)$ which has properties 3—6 and 8, but not properties 7, 9 or 10.

Example 5[25]). Let M_0 denote the indecomposable continuum described by Knaster in his thesis[26]), and denoted there by K_1. It lies

[24]) See pp. 75—76 of the paper of footnote [23]).

[25]) Examples 1—5 are Gehman's examples 1, 2, 3, 8 and 10.

[26]) B. Knaster, Un continu dont tous sous-continu est indécomposable, Fund. Math. **3** (1922), pp. 246—286. This example is described on pp. 269—271.

within the square whose vertices are $(0, 0)$, $(0, 1)$, $(1, 1)$, $(1, 0)$ and is irreducible between the point $(1, 0)$ and any point that cannot be joined to $(1, 0)$ by an arc in M_0. Let M_n $(n = 1, 2, \ldots)$ be the set of points (x', y') obtained from the points (x, y) of M_0 by the transformation:

$$x' = (x + 2^{n+1} - 2)/2^n, \quad y' = y/2^n.$$

Let p be the point $(2, 0)$, and let

$$M = p + \sum_{i=0}^{\infty} M_i.$$

The point p has properties 3—6 and 8—10, but not property 7.

Example 6. Let M consist of the intervals from $(2, 0)$ to $(0, 0)$, from $(1/n, 0)$ to $(1/n, 1/n)$, from $(1/n, 1/n)$ to $(-1/n, 1/n)$, from $(-1/n, 1/n)$ to $(-1/n, -1/n)$, from $(-1/n, -1/n)$ to $(2, 0)$, for $n = 1, 2, \ldots$. Let p be the point $(0, 0)$, which has properties 4, 5, 8, and 10, but not 3, 6, 7, and 9.

Example 7. In euclidean space of three dimensions take a cylindrical coördinate system (θ, r, z). Let M_0 consist of the intervals from $(0, 0, 0)$ to $(0, 0, 1)$, and from $(0, 0, 1)$ to $(0, 1, 1)$. Let M_1 consist of the intervals from $(\pi/2^i, 0, 1/2^i)$ to $(\pi/2^i, 1/2^i, 1/2^i)$, from $(\pi/2^i, 1/2^i, 1/2^i)$ to $(\pi/2^i, 1/2^i, 1)$, from $(\pi/2^i, 1/2^i, 1)$ to $(\pi/2^i, 3/4, 1)$, for $i = 1, 2, \ldots$. Let M_{n+1} $(n = 1, 2, \ldots)$ be the set of all points (θ', r', z'), where (θ, r, z) is a point of M_n,

$$\theta' = (\theta + \pi)/2^k, \quad r' = r/2^k, \quad z' = z/2^k,$$

and $k = 1, 2, \ldots$. Let

$$M = \sum_{n=0}^{\infty} M_n,$$

and let p be the point $(0, 0, 0)$. The point p has properties 4—6, 8, and 10, but not properties 3, 7, and 9.

That M does not have property 3 at p may be seen as follows: Let $N = M_0$. We will show that $M - N + p$ is connected. Suppose

$$M - N + p = K_1 + K_2, \quad \bar{K}_1 \cdot K_2 + K_1 \cdot \bar{K}_2 = 0, \quad K_1 \supset p.$$

Consider the first broken line $(i = 1)$ of M_1. The sequence of broken lines of M_2 for $k = 1$ converges down on the interval $(\pi/2, 0, 1/2)$ to $(\pi/2, 1/2, 1/2)$ of this first broken line of M_1. Hence all but a finite number of this sequence of M_2 must belong to the same set K_i as the line of M_1. But p is a limit point of this sequence of M_2 and all but a finite number of the sequence must belong to K_1. Therefore the first broken line of M_1 belongs to K_1 (except for the point $(\pi/2, 0, 1/2)$). Similarly if we consider *any* broken line of M_n, we find M_{n+1} contains a sequence of broken lines converging on p and an interval of the line of M_n. Thus every one

of the broken lines must belong to K_1 and $K_2 = 0$. Hence $M - N + p$ is connected and p does not have property 3. However p does have property 6 as each of the components of $M - N$ is a single broken line and no one of them has p as a limit point.

15. In Gehman's paper the question as to whether property 3 is stronger than property 6 is left unsolved. In example 7 we have a point which has property 6 but not property 3, and thus property 3 is stronger than 6. Gehman also raised the question as to whether examples exist of a point having properties 4—6 and 8 but not 3, 7 and 9, and of a point having properties 4—6 but not 3 and 7—9. In the first case the point p of example 7 has exactly the properties desired. Now let M^* be the set M of example 7 plus the set of semi-circles with centers $(0, 0, 3/2^{n+3})$ and radii $1/2^{n+3}$ lying on one side of the z-axis in the plane $\theta = 0$. Then the point p of M^* has properties 4—6, but not 3 and 7—9. This is the second desired example.

Gehman[27]) proved that the set of all points having property 3 is totally disconnected and the set of all points having property 6 contains no continuum. The question as to whether the set of all points having property 6 is totally disconnected is not discussed. We shall settle this question in the following

Theorem 21. *The set of all points of a continuum which have property* 6 *is totally disconnected.*

If the theorem were not true, there would exist a continuum M containing a connected subset H of more than one point and such that every point of H has property 6. If $H = M$, let x and y be points of M and N be an irreducible subcontinuum of M between x and y. Every point of $N - x - y$ belongs to H and has property 6. As property 6 implies property 5, no point of H belongs to a subcontinuum of M which is irreducible between two other points. But obviously every point of $N - x - y$ belongs to such a continuum, and thus $H \neq M$. Let $p \in H$ and $q \in M - H$.

If $H - p$ contains a connected subset G having p as a limit point, let K be a subcontinuum of M which is irreducible between p and q. As above $H(K - p - q) = 0$. Then $K \cdot G = 0$, and p is a limit point of the connected subset G of $M - K$. This is contrary to our assumption that p has property 6.

If $H - p$ contains no connected subset having p as a limit point, let G be a component of $H - p$. By a theorem due to Knaster and

[27]) See Theorem 19 and Corollary 21a of the paper of [23]).

Kuratowski[28]) the set $H - G$ is connected. Since G is a component of $H - p$ and does not have p as a limit point, G is closed relative to H. Then, as H is connected, $H - G$ has a limit point x in G. Let K be a subcontinuum of M which is irreducible between q and x. As every point of H has property 5, $H(K - q - x) = 0$. Hence $K(H - G) = 0$ and x is a limit point of the connected subset $H - G$ of $M - K$. Then x does not have property 6. But this is a contradiction as every point of H has property 6.

16. In connection with these end point properties there remain as yet several interesting unsolved problems:

(a) In a *plane* continuum are any of the following relations true: property 6 implies 3, 6 plus 8 implies 3, 6 plus 10 implies 3, 6 plus 8 plus 10 implies 3? Example 7 shows that none of these relations exist in general; but as the continuum here is homeomorphic with no plane continuum, the questions remain open for a plane continuum.

(b) Is the set of all points of a continuum having property 5 totally disconnected? Either it is totally disconnected or there exists a connected set H such that if x and y are points of $\overline{H}$ and K is a subcontinuum of $\overline{H}$ irreducible between x and y, then $K - x - y \subset \overline{H} - H$.

(c) Is the set of all points having property 10 totally disconnected?

(d) For what properties is the set of all points having that property zero-dimensional? The set of all points having property 7 is zero-dimensional. The set of all points having property 4, or 8, or 9 may contain a continuum, and thus may not be zero-dimensional. But what about properties 3, 6, 5, 10, and 5 plus 9?

---

[28]) B. Knaster and C. Kuratowski, Sur les ensembles connexes, Fund. Math. 2 (1921), pp. 206—255, theorem X.

(Eingegangen am 28. Oktober 1929.)

Sätze über algebraische Ringe.

Von

Trygve Nagell in Oslo.

Es sei gegeben der algebraische Zahlkörper K vom Grade n. Man verstehe unter einem *Ring* jeden Integritätsbereich von ganzen Körperzahlen. Wir betrachten nur diejenigen Ringe, die Zahlen n-ten Grades enthalten. Genau wie für Körper beweist man, daß solche Ringe eine Basis von n Zahlen besitzen. Es sei ξ eine Zahl n-ten Grades in dem Ringe R und, wenn $\Omega_1, \Omega_2, \ldots, \Omega_n$ eine Körperbasis bedeuten,

$$\xi^k = c_{k1}\Omega_1 + c_{k2}\Omega_2 + \ldots + c_{kn}\Omega_n \qquad (0 \leqq k < n),$$

mit ganzen rationalen c. Durch Umkehrung folgt hieraus

$$\Omega_m = \frac{1}{|c_{kj}|}\cdot(b_{m1} + b_{m2}\xi + \ldots + b_{mn}\xi^{n-1}) \quad (1 \leqq m \leqq n),$$

wo die b ganz rational sind, und $|c_{kj}|$ die Determinante von den Koeffizienten c bedeutet. Es sei C der absolute Wert von $|c_{kj}|$. Dann gilt

Satz I. *Für jedes i der Reihe $i = 1, 2, \ldots, n$ sei a_{ii} die kleinste natürliche Zahl, so daß Zahlen von der Form*

$$\omega_i = \frac{a_{i1} + a_{i2}\xi + \ldots + a_{ii}\xi^{i-1}}{C} \tag{1}$$

im Ringe R vorhanden sind, mit sonst beliebigen ganzen rationalen Koeffizienten a. Dann bildet jedes System $\omega_1, \omega_2, \ldots, \omega_n$ eine Basis des Ringes.

Beweis. Jede Zahl α des Ringes hat die Form

$$\alpha = \frac{b_1 + b_2\xi + \ldots + b_n\xi^{n-1}}{C}$$

mit ganzen rationalen b. Es werde b_n durch a_{nn} dividiert:

$$b_n = h_n a_{nn} + r_n, \qquad 0 \leqq r_n < a_{nn}.$$

Dann gehört auch die Zahl

$$\alpha - h_n \omega_n = \frac{b_1' + b_2' \xi + \ldots + b_{n-1}' \xi^{n-2} + r_n \xi^{n-1}}{C}$$

zu dem Ring; wegen der Minimaleigenschaft von a_{nn} ist notwendig $r_n = 0$. Auf dieselbe Weise beweist man, daß b_{n-1}' durch $a_{n-1,\,n-1}$ teilbar ist, und folglich

$$\alpha - h_n \omega_n - h_{n-1} \omega_{n-1} = \frac{b_1'' + b_2'' \xi + \ldots + b_{n-2}'' \xi^{n-3}}{C}.$$

So geht es weiter bis zu

$$\alpha - h_n \omega_n - h_{n-1} \omega_{n-1} - \ldots - h_1 \omega_1 = 0.$$

Die Diskriminante des Ringes ist gegeben durch

$$D(\omega_1, \omega_2, \ldots, \omega_n) = \frac{(a_{11} a_{22} \ldots a_{nn})^2}{C^{2n}} \cdot D(\xi). \tag{2}$$

Für die durch Satz I gegebene Basis gilt ferner

Satz II. *Das Quadrat der ganzen rationalen Zahl* $\omega_1 = \frac{a_{11}}{C}$ *ist ein Teiler von* $D(\omega_1, \omega_2, \ldots, \omega_n)$.[1]) *Die Zahlen* a_{ii} $(i > 1)$ *sind Teiler von* C. *Es ist immer* a_{ii} *durch* a_{jj} *teilbar, wenn* $j \geqq i$ *ist. Die Zahlen* a_{ik} $(i \neq k)$ *sind alle durch* a_{ii} *teilbar, und sie können so gewählt werden, daß*

$$\begin{cases} 0 \leqq a_{ik} < C & (i > k > 1), \\ 0 \leqq a_{i1} < a_{11} = \omega_1 C. \end{cases} \tag{3}$$

Beweis. Da $D(1, \omega_2, \ldots, \omega_n)$ eine ganze rationale Zahl ist, muß die Diskriminante des Ringes durch ω_1^2 teilbar sein. Es gibt ganze rationale Zahlen x, so daß für alle $i > 1$

$$\xi^{i-1} = x_1 \omega_1 + x_2 \omega_2 + \ldots + x_i \omega_i = \frac{x_i a_{ii}}{C} \xi^{i-1}.$$

Hieraus folgt, daß a_{ii} in C aufgehen muß. Es ist ferner

$$\omega_i \xi^{j-i} = \frac{a_{i1} \xi^{j-i} + \ldots + a_{ii} \xi^{j-1}}{C} \qquad (j \geqq i)$$

Folglich ist a_{ii} ein Multiplum von a_{jj} für alle $j \geqq i$.

Daß a_{ik} durch a_{ii} teilbar ist, wollen wir durch Induktion beweisen. Für $i = 2$ ist die Behauptung richtig; denn es ist offenbar

$$\xi = \omega_2 \frac{C}{a_{22}} + \frac{a_{21}}{a_{22}},$$

---

[1]) Es gibt von Null verschiedene, ganze rationale Zahlen im Ring; denn $N(\xi$ gehört offenbar zum Ring. ω_1 ist die kleinste natürliche Zahl des Ringes.

nd da C durch a_{22} teilbar ist, muß es folglich auch a_{21} sein. Es sei nun ngenommen, daß die Behauptung für $i = 2, 3, \ldots, j$ richtig ist.

Es gibt ganze rationale Zahlen x, so daß

$$\omega_j \xi = \frac{a_{j1}\xi + a_{j2}\xi^2 + \ldots + a_{jj}\xi^j}{C} = x_1\omega_1 + x_2\omega_2 + \ldots + x_{j+1}\omega_{j+1}.$$

Hieraus ergibt sich

$$x_{j+1} = \frac{a_{jj}}{a_{j+1,\,j+1}},$$

nd für alle $k \leqq j$

$$x_k a_{kk} + x_{k+1} a_{k+1,k} + \ldots + x_j a_{jk} + x_{j+1} a_{j+1,k} = a_{j,k-1},$$

olglich

$$\frac{a_{j+1,k}}{a_{j+1,j+1}} = \frac{a_{j,k-1}}{a_{jj}} - \sum_{m=k}^{m=j} x_m \cdot \frac{a_{mk}}{a_{jj}}. \tag{4}$$

Die Zahlen a_{mk} sind alle durch a_{mm} teilbar, unserer Voraussetzung ufolge. Wegen $m \leqq j$ ist a_{mm} durch a_{jj} teilbar. Die Zahlen $\frac{a_{mk}}{a_{jj}}$ sind olglich ganz, wenn $k \leqq m \leqq j$ ist. Aus der Gleichung (4) folgt dann, daß $_{j+1,k}$ durch $a_{j+1,\,j+1}$ teilbar ist. Unsere Behauptung gilt also auch für $= j+1$ und folglich für alle i.

Da man offenbar die Zahlen a_{ik} $(2 \leqq k < i)$ modulo C und die Zahlen a_{i1} $> 1)$ modulo $\omega_1 C$ reduzieren kann, ist auch die letzte Behauptung des atzes II richtig.

Es sei nun $\Omega_1, \Omega_2, \ldots, \Omega_n$ eine Körperbasis und

$$\omega_i = b_{i1}\Omega_1 + b_{i2}\Omega_2 + \ldots + b_{in}\Omega_n \qquad (1 \leqq i \leqq n),$$

it ganzen rationalen b. Durch Umkehrung folgt hieraus

$$\Omega_m = \frac{1}{|b_{ij}|} \cdot (g_{m1}\omega_1 + g_{m2}\omega_2 + \ldots + g_{mn}\omega_n) \quad (1 \leqq m \leqq n),$$

o die g ganz rational sind, und wo $|b_{ij}|$ die Determinante der Koeffienten b bedeutet. Es gilt folglich: *Es sei α eine beliebige ganze Körperahl und A eine beliebige natürliche Zahl. Gibt es dann im Körper Ringe it der Diskriminante $A^2 \cdot D(\Omega_1, \Omega_2, \ldots, \Omega_n)$, so gehört die Zahl αA allen diesen Ringen.*

Hieraus folgt nun weiter

Satz III. *Es gibt im Körper nur endlich viele Ringe mit gegebener iskriminante D.*

Beweis. Wenn α eine ganze Körperzahl n-ten Grades ist, und wenn $= A^2 \cdot D(\Omega_1, \Omega_2, \ldots, \Omega_n)$ ist, setzen wir $\xi = \alpha A$. In jedem Ring mit r Diskriminante D gibt es dann eine mit Hilfe von ξ gebildete Basis

wie im Satz I angegeben, die den im Satz II genannten Nebenbedingungen genügt. Für eine solche Basis bestehen aber nur endlich viele Möglichkeiten. Dieser Beweis liefert zugleich eine Methode, um zu entscheiden, ob Ringe mit der gegebenen Diskriminante existieren, und um alle eventuell existierenden zu bestimmen.

Aus den Sätzen I und II folgt ferner sofort: *Eine Zahl n-ten Grades kann nur in endlich vielen Ringen des Körpers vorkommen.*

Wenn der Ring die Zahl 1 enthält, können wir offenbar in den Sätzen I und II die Zahl C durch die Zahl $\sqrt{\frac{D(\xi)}{D(\omega_1, \omega_2, \ldots, \omega_n)}}$ ersetzen.

(Eingegangen am 1. Februar 1930.)

Zur algebraischen Zahlentheorie.

Von

Trygve Nagell in Oslo.

Einleitung.

1. Gibt uns die allgemeine Theorie der algebraischen Zahlen und der Ideale zugleich die Mittel, um in einem willkürlich gegebenen Zahlkörper die verschiedenen numerischen Aufgaben zu lösen, die uns gestellt werden? Können wir z. B. immer eine Basis der ganzen Zahlen des Körpers berechnen? Können wir ein System von Fundamentaleinheiten bestimmen? Können wir entscheiden, ob ein vorgelegtes Ideal ein Hauptideal ist oder nicht? Um diese und ähnliche Fragen haben sich die Zahlentheoretiker bisher sehr wenig gekümmert. Die üblichen Beweise in der allgemeinen Theorie sind nur Existenzbeweise, die nicht zugleich ein Verfahren geben, um in jedem einzelnen numerischen Fall die Größen zu berechnen, deren Existenz bewiesen wird. Die Theorie kann aber nicht als vollständig angesehen werden, bevor diese Fragen gelöst sind. Bisher war dies nur der Fall für die quadratischen Körper. Ich zeige in dieser Arbeit, wie man — im allgemeinen durch kleine Veränderungen der Beweise — die wichtigsten numerischen Aufgaben in jedem beliebigen algebraischen Zahlkörper lösen kann. Ob die angegebenen Methoden die bestmöglichen sind, steht dahin.

In der Folge sei gegeben der algebraische Zahlkörper $K(\xi)$ n-ten Grades, erzeugt durch die Zahl ξ, eine auf irgendeine Weise charakterisierte Wurzel der Gleichung

$$\xi^n + a_1 \xi^{n-1} + \ldots + a_{n-1}\xi + a_n = 0 .$$

Es wird vorausgesetzt, daß man die Wurzeln dieser Gleichung mit beliebiger Genauigkeit numerisch berechnen kann. Wir können ξ ganz annehmen. Die zu einer Zahl α des Körpers konjugierten Zahlen werden wie üblich mit

$$\alpha = \alpha^{(1)}, \alpha^{(2)}, \ldots, \alpha^{(n)}$$

bezeichnet

§ 1.

Körperbasis, Körperdiskriminante usw.

2. Die Norm der Zahl ξ ist gegeben durch $N(\xi) = (-1)^n a_n$. Die Diskriminante $D(\xi)$ ist gleich der Diskriminante der Gleichung, die man mittels der Formeln der Algebra berechnet. Zur Berechnung einer Basis der ganzen Zahlen des Körpers dient der folgende Satz:

Für jedes i der Reihe $i = 2, 3, \ldots, n$ sei a_{ii} die kleinste natürliche Zahl, so daß ganze Zahlen von der Form

$$\omega_i = \frac{a_{i1} + a_{i2}\xi + \ldots + a_{ii}\xi^{i-1}}{D(\xi)} \tag{1}$$

vorhanden sind, mit sonst beliebigen ganzen rationalen Koeffizienten a. Dann bildet jedes System $\omega_1 = 1, \omega_2, \ldots, \omega_n$ eine Basis der ganzen Zahlen des Körpers.

Beweis. Jede ganze Zahl α hat die Form

$$\alpha = \frac{b_1 + b_2\xi + \ldots + b_n\xi^{n-1}}{D(\xi)}$$

mit ganzen rationalen b. Es werde b_n durch a_{nn} dividiert

$$b_n = h_n a_{nn} + r_n, \qquad 0 \leqq r_n < a_{nn}.$$

Dann ist auch die Zahl

$$\alpha - h_n\omega_n = \frac{b_1' + b_2'\xi + \ldots + b_{n-1}'\xi^{n-2} + r_n\xi^{n-1}}{D(\xi)}$$

ganz; wegen der Minimaleigenschaft von a_{nn} ist notwendig $r_n = 0$. Auf dieselbe Weise beweist man, daß b_{n-1}' durch $a_{n-1,n-1}$ teilbar ist, und folglich

$$\alpha - h_n\omega_n - h_{n-1}\omega_{n-1} = \frac{b_1'' + b_2''\xi + \ldots + b_{n-2}''\xi^{n-3}}{D(\xi)}.$$

So geht es weiter bis zu

$$\alpha - h_n\omega_n - h_{n-1}\omega_{n-1} - \ldots - h_1\omega_1 = 0.$$

Wir haben offenbar für jedes i

$$0 < a_{ii} \leqq |D(\xi)|. \tag{2}$$

Die übrigen Koeffizienten können wir so wählen, daß

$$0 \leqq a_{ik} < |D(\xi)| \qquad (i + k). \tag{2'}$$

Die Aufgabe, eine Basis zu bestimmen, ist also darauf reduziert, zu untersuchen, welche von den (endlich vielen) Zahlen (1) mit den Nebenbedingungen (2) und (2') ganz sind. Damit die Zahl ω_i ganz ist, ist notwendig und hinreichend, daß die symmetrischen Grundfunktionen von ω_i und ihren

Konjugierten $\omega_i^{(2)}, \ldots, \omega_i^{(n)}$ ganze rationale Zahlen sind. Es ist übrigens leicht folgendes zu beweisen: Die Zahlen a_{ii} müssen alle in $D(\xi)$ aufgehen. Es ist immer a_{ii} durch a_{jj} teilbar, wenn $i \leqq j$ ist. Die Zahlen a_{ik} sind alle durch a_{ii} teilbar. Wenn $D(\xi) = AB^2$ ist, mit quadratfreiem A, kann man hier sogar überall $D(\xi)$ durch B ersetzen.

Die Körperdiskriminante ist gegeben durch die Gleichung

$$D(\omega_1, \omega_2, \ldots, \omega_n) = \frac{(a_{22}\, a_{33} \ldots a_{nn})^2}{D(\xi)^{2n-3}}.$$

In der Folge bedeutet überall $\omega_1, \omega_2, \ldots, \omega_n$ eine Körperbasis.

3. Es sei gegeben die Zahl

$$\alpha = x_1 + x_2 \xi + \ldots + x_n \xi^{n-1}$$

mit rationalen Koeffizienten x. Man soll entscheiden, von welchem Grade α ist. Dann berechnet man die symmetrischen Grundfunktionen $c_1, c_2, \ldots, c_n$ von α und ihren Konjugierten. Ist dann identisch

$$x^n + c_1 x^{n-1} + \ldots + c_n = [g(x)]^{\frac{n}{m}},$$

wo $g(x)$ ein Polynom m-ten Grades ohne mehrfache Wurzeln ist, so ist α vom m-ten Grade. Dies kann durch eine endliche Anzahl von Versuchen entschieden werden. Es ist $N(\alpha) = (-1)^n c_n$.

Aus der vorigen Nummer folgt, wie man die Koeffizienten r in der Gleichung

$$\alpha = r_1 \omega_1 + r_2 \omega_2 + \ldots + r_n \omega_n$$

bestimmt.

Es seien gegeben die n positiven Zahlen $g_1, g_2, \ldots, g_n$. Dann gibt es nur endlich viele ganze Körperzahlen $\alpha = \alpha^{(1)}$, so daß für alle i der Reihe $i = 1, 2, \ldots, n_1$,

$$|\alpha^{(i)}| < g_i.$$

Diese bestimmt man am einfachsten so: Man setzt

$$\alpha^{(i)} = x_1 \omega_1^{(i)} + x_2 \omega_2^{(i)} + \ldots + x_n \omega_n^{(i)} \qquad (1 \leqq i \leqq n)$$

und löst diese Gleichungen in bezug auf x_m. Dann ergibt sich für x_m ein Ausdruck

$$x_m = \frac{d_m}{\sqrt{D(\omega_1, \omega_2, \ldots, \omega_n)}} \qquad (1 \leqq m \leqq n),$$

wo d_m diejenige Determinante bedeutet, die man erhält, wenn man in der Determinante $|\omega_k^{(i)}|$ die Zahlen $\omega_m^{(i)}$ durch die Zahlen $\alpha^{(i)}$ ersetzt. Wenn die Körperbasis bekannt ist, folgt aus dieser Gleichung, wegen $|\alpha^{(i)}| < g_i$, eine obere Grenze für den absoluten Wert der ganzen rationalen Zahl x_m. Durch eine endliche Anzahl von Versuchen kann man dann die Zahlen α bestimmen.

4. Es sei gegeben eine algebraische Zahl β, Wurzel der Gleichung

$$\beta^m + b_1\beta^{m-1} + \ldots + b_m = 0, \tag{3}$$

mit rationalen Koeffizienten b, und es soll entschieden werden, ob sie zu dem Körper $K(\xi)$ gehört oder nicht. Es wird angenommen, daß β so charakterisiert ist, daß man sie mit beliebiger Genauigkeit numerisch berechnen kann. Dann setzen wir

$$\beta = x_1 + x_2\xi + \ldots + x_n\xi^{n-1}, \tag{4}$$

mit den n Unbestimmten $x_1, x_2, \ldots, x_n$. Wird dieser Wert in die Gleichung (3) eingetragen, so erhalten wir eine Gleichung

$$R_1(x)\xi^{n-1} + R_2(x)\xi^{n-2} + \ldots + R_n(x) = 0,$$

wo die $R_i(x)$ ganze rationale Funktionen der Unbestimmten $x_1, x_2, \ldots, x_n$ sind, mit rationalen Koeffizienten. Wir müssen dann entscheiden, ob das simultane System

$$R_1(x) = 0,\ R_2(x) = 0,\ \ldots,\ R_n(x) = 0$$

rationale Lösungen hat, und endlich, ob diese Werte von $x_1, x_2, \ldots, x_n$ in (4) eingetragen wirklich β liefern.

Die Aufgabe kann aber auch auf die folgende Weise gelöst werden: β können wir offenbar ganz annehmen. Aus der Gleichung (3) folgen nun leicht Abschätzungen für β und ihre Konjugierten (in einem beliebigen Zahlkörper). Es ist z. B.

$$|\beta^{(i)}| \leqq |b_1| + |b_2| + \ldots + |b_m|.$$

Wenn nun β zu dem Körper $K(\xi)$ gehört, und wenn

$$\beta = x_1\omega_1 + x_2\omega_2 + \ldots + x_n\omega_n$$

ist, bestehen nur endlich viele Möglichkeiten für die ganzzahligen Koeffizienten x, die nach der vorigen Nummer bestimmt werden können, wenn die Körperbasis bekannt ist. Durch endlich viele Versuche kann dann entschieden werden, ob β zu dem Körper gehört oder nicht.

Um die Aufgabe vollständig zu lösen, bleibt es aber noch übrig zu entscheiden, ob es unter den gefundenen Zahlen

$$\eta = x_1 + x_2\xi + x_3\xi^2 + \ldots + x_n\xi^{n-1},$$

eventuell

$$\eta = x_1\omega_1 + x_2\omega_2 + \ldots + x_n\omega_n,$$

wirklich eine mit β identische Zahl gibt. Nur solche Zahlen η können in Betracht kommen, die Wurzeln der Gleichung

$$x^m + b_1x^{m-1} + b_2x^{m-2} + \ldots + b_m = 0$$

sind. Ohne Einschränkung der Allgemeinheit können wir annehmen, daß die Diskriminante D dieser Gleichung von Null verschieden ist. Wenn $\eta_1, \eta_2, \ldots, \eta_m$ die Wurzeln der Gleichung bedeuten, folgt dann leicht

$$|\eta_i - \eta_j| > \varphi = \frac{|\sqrt{D}|}{(2|b_1|+2|b_2|+\ldots+2|b_m|)^{\frac{1}{2}m(m-1)-1}} \quad (i \neq j).$$

Wir bestimmen nun eine positive ganze Zahl M, die größer als $\frac{3\sqrt{2}}{\varphi}$ ist. Ist dann (mit reellen Größen A und B)

$$M\eta_i = A_i + B_i\sqrt{-1} \quad \text{und} \quad M\eta_j = A_j + B_j\sqrt{-1},$$

so folgt, daß wenigstens eine der Zahlen

$$|A_i - A_j| \quad \text{und} \quad |B_i - B_j|$$

größer als 3 ist. Unsere Frage können wir nun offenbar wie folgt entscheiden: Wenn $M\eta = A + B\sqrt{-1}$ und $M\beta = C + D\sqrt{-1}$ ist (A, B, C und D reell), so braucht man von A, B, C und D nur die Ziffern vor dem Dezimalkomma zu berechnen. (Diese Ziffern sind in gewissen Fällen — z. B. wenn A eine ganze rationale Zahl ist — nur bis auf eine Einheit bestimmt.) Wenn die Ziffern (und das Vorzeichen) von A und C übereinstimmen, eventuell bis auf eine Einheit, und wenn dasselbe für B und D gilt, dann ist $\eta = \beta$ und nur dann.

§ 2.

Einheiten.

5. Bestimmung der Einheitswurzeln des Körpers. Man bestimmt die sämtlichen Einheitswurzeln, deren Grad ein Teiler von n ist, und entscheidet, welche von diesen zu dem Körper gehören (vgl. Nummer 4).

Bestimmung von r unabhängigen Einheiten[1]). Es gibt offenbar (mindestens) ein System von r unabhängigen Einheiten $\eta_1, \eta_2, \ldots, \eta_r$, so daß die Determinante der Logarithmen

$$d = |\log|\eta_k^{(i)}|| \qquad \begin{pmatrix} 1 \leqq i \leqq r \\ 1 \leqq k \leqq r \end{pmatrix}$$

absolut genommen >1 ist. Denn ist $|d| \leqq 1$ und bezeichnet N die größte ganze Zahl $\leqq \frac{1}{|d|} + 1$, so hat das System $\eta_1^N, \eta_2, \ldots, \eta_r$, die Determinante $d_1 = Nd$. Es ist aber offenbar $|Nd| > 1$.

---

[1]) Es ist $r = r_1 + r_2 - 1$, wo r_1 die Anzahl der reellen und $2r_2$ die Anzahl der imaginären von den konjugierten Körpern bedeutet.

Hieraus folgt nun die folgende Methode zur Bestimmung von r unabhängigen Einheiten. Wenn $\omega_1, \omega_2, \ldots, \omega_n$ eine Körperbasis bedeutet, betrachten wir die Zahlen

$$x_1\omega_1 + x_2\omega_2 + \ldots + x_n\omega_n,$$

für die die absoluten Werte der Koeffizienten $x_1, x_2, \ldots, x_n$ höchstens gleich m sind, und greifen diejenigen Zahlen heraus, deren Norm gleich ± 1 ist. Wird m allmählich größer und größer gewählt, so müssen wir nach einer endlichen Anzahl von Versuchen auf r Einheiten $\eta_1, \eta_2, \ldots, \eta_r$ stoßen, mit der folgenden Eigenschaft: In der (gewöhnlichen) Dezimalbruchentwicklung von dem absoluten Wert der Determinante der Logarithmen

$$d = \left|\log\left|\eta_k^{(i)}\right|\right| \qquad \begin{pmatrix} 1 \leqq i \leqq r \\ 1 \leqq k \leqq r \end{pmatrix}$$

steht links vom Komma eine ganze Zahl $\geqq 1$. Dann ist ja $d \neq 0$.

Bestimmung eines Systems von Fundamentaleinheiten. Wir benutzen E. Landau, Vorlesungen über Zahlentheorie (Leipzig 1927), Bd. III. Die auf Seite 166 dieses Werkes erwähnten Einheiten $H_1, H_2, \ldots, H_u$ können offenbar bestimmt werden, wenn das System von unabhängigen Einheiten $\eta_1, \eta_2, \ldots, \eta_r$ gegeben ist. Die Zahl j ist dann offenbar ein Teiler von $u! = M$. Für jede Einheit ε des Körpers besteht dann eine Gleichung

$$\varepsilon^M = \eta_1^{t_1}\eta_2^{t_2}\ldots\eta_r^{t_r}$$

mit ganzen rationalen t. Hieraus folgt der Satz (siehe S. 167):

Für jedes m der Reihe $1, 2, 3, \ldots, r$ sei t_{mm} die *kleinste* natürliche Zahl, so daß Einheiten von der Form

$$\varepsilon_m = \sqrt[M]{\eta_1^{t_{m1}}\eta_2^{t_{m2}}\ldots\eta_m^{t_{mm}}} \tag{1}$$

im Körper vorhanden sind, mit sonst beliebigen ganzen rationalen Exponenten t. Dann bilden die Zahlen $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_r$ ein System von Fundamentaleinheiten.

Wir haben hier offenbar für jedes m

$$0 < t_{mm} \leqq M. \tag{2}$$

Die übrigen Exponenten können wir so wählen, daß

$$0 \leqq t_{mk} < M \qquad (m \neq k). \tag{2'}$$

Um unsere Aufgabe zu lösen, haben wir nun nur zu untersuchen, welche von den (endlich vielen) Zahlen (1) mit den Nebenbedingungen (2) und (2′) zu dem Körper gehören (vgl. Nummer 4). Die gesuchten Zahlen ε_m sind diejenigen mit einem Minimum von t_{mm}.

§ 3.

Ideale.

6. Es sei gegeben das Ideal

$$\mathfrak{j} = (\beta_1, \beta_2, \ldots, \beta_l),$$

und es soll entschieden werden, ob die ganze Zahl

$$\alpha = a_1 \omega_1 + a_2 \omega_2 + \ldots + a_n \omega_n$$

zu dem Ideal gehört. Dann muß eine Gleichung

$$\alpha = \lambda_1 \beta_1 + \lambda_2 \beta_2 + \ldots + \lambda_l \beta_l$$

bestehen, mit ganzen λ. Es sei nun

$$\beta_i = b_{i1} \omega_1 + b_{i2} \omega_2 + \ldots + b_{in} \omega_n \qquad (1 \leqq i \leqq l)$$

und ferner mit unbestimmten Koeffizienten x

$$\lambda_i = x_{i1} \omega_1 + x_{i2} \omega_2 + \ldots + x_{in} \omega_n \qquad (1 \leqq i \leqq l).$$

Durch Einführung dieser Werte in die Gleichung für α und Vergleichung der Koeffizienten von ω_k erhält man hieraus n lineare Gleichungen mit nl Unbestimmten

$$a_k = \sum_{i=1}^{l} \sum_{j=1}^{n} c_{ijk} x_{ij} \qquad (1 \leqq k \leqq n),$$

mit ganzen rationalen c. Aus der elementaren Zahlentheorie ist bekannt, wie man entscheidet, ob dieses simultane System Lösungen in ganzen rationalen x hat.

Zwei Ideale $\mathfrak{j} = (\beta_1, \beta_2, \ldots, \beta_l)$ und $\mathfrak{j}_1 = (\gamma_1, \gamma_2, \ldots, \gamma_m)$ sind dann und nur dann identisch, wenn jede der Zahlen γ zu $\mathfrak{j}$ gehört, und wenn jede der Zahlen β zu $\mathfrak{j}_1$ gehört. Dies kann immer entschieden werden.

7. Bestimmung einer Basis des Ideals. Es sei gegeben dasselbe Ideal $\mathfrak{j}$ wie in der vorigen Nummer. Dann gilt der Satz[2]):

Für jedes m der Reihe $m = 1, 2, \ldots, n$ sei a_{mm} die kleinste natürliche Zahl, so daß im Ideal Zahlen von der Form

$$\alpha_m = a_{m1} \omega_1 + a_{m2} \omega_2 + \ldots + a_{mm} \omega_m \qquad (1)$$

vorhanden sind, mit sonst beliebigen ganzen rationalen Koeffizienten a. Dann bilden die Zahlen $\alpha_1, \alpha_2, \ldots, \alpha_n$ eine Basis des Ideals.

Es sei nun a die kleinste natürliche Zahl des Ideals. Diese kann offenbar durch eine endliche Anzahl von Versuchen bestimmt werden. Die Zahlen $a_{11}, a_{22}, \ldots, a_{nn}$ müssen dann alle Teiler von a sein. Die übrigen

---

[2]) Siehe z. B. E. Landau, Vorles. über Zahlentheorie (Leipzig 1927), Bd. III, S. 117–118.

Koeffizienten können wir so wählen, daß

$$0 \leqq a_{mk} < a \qquad (m \neq k)$$

Die Aufgabe, eine Basis zu bestimmen, ist also darauf reduziert, zu entscheiden, welche von den (endlich vielen) Zahlen (1) mit den soeben angeführten Nebenbedingungen zu dem Ideale gehören. Statt a hätten wir auch den absoluten Wert der Norm einer beliebigen (von Null verschiedenen) Zahl des Ideals nehmen können. Wenn $\omega_1 = 1$, so ist $a_{11} = a$. Die Norm des Ideals ist

$$a_{11} a_{22} \dots a_{nn}.$$

Es sei gegeben das Ideal $\mathfrak{a}$ mit der Basis $\alpha_1, \alpha_2, \dots, \alpha_n$, und das Ideal $\mathfrak{b}$ mit der Basis $\beta_1, \beta_2, \dots, \beta_n$. Wir bestimmen die rationalen Zahlen a, so daß

$$\beta_i = a_{i1}\alpha_1 + a_{i2}\alpha_2 + \dots a_{in}\alpha_n \qquad (1 \leqq i \leqq n).$$

Damit die Ideale $\mathfrak{a}$ und $\mathfrak{b}$ identisch sind, ist notwendig und hinreichend, daß die Koeffizienten a ganz sind und daß die Determinante $|a_{ij}| = \pm 1$ ist.

8. Bestimmung aller Ideale mit gegebener Norm. Ein Ideal $\mathfrak{j}$ mit der Norm N muß (wenn wir in der vorigen Nummer $\omega_1 = 1$ wählen) eine Basis

$$(2) \qquad \begin{cases} \alpha_1 = a_{11}, \\ \alpha_2 = a_{21} + a_{22}\omega_2, \\ \dots\dots\dots\dots\dots\dots \\ \alpha_n = a_{n1} + a_{n2}\omega_2 + \dots + a_{nn}\omega_n, \end{cases}$$

mit den folgenden Eigenschaften besitzen: Es bedeutet a_{mm} die kleinste natürliche Zahl, so daß im Ideal eine Zahl

$$\alpha_m = a_{m1} + a_{m2}\omega_2 + \dots + a_{mm}\omega_m \qquad (1 \leqq m \leqq n)$$

vorhanden ist. Die Zahlen a_{mm} $(2 \leqq m \leqq n)$ sind alle Teiler von a_{11}. Es ist ferner

$$(3) \qquad N = a_{11} a_{22} \dots a_{nn}$$

und

$$0 \leqq a_{mk} < a_{11} \qquad (m \neq k).$$

Wir berechnen nun die sämtlichen Systeme (2) mit diesen Nebenbedingungen. Alle Ideale mit der Norm N müssen sich dann unter den Idealen

$$\mathfrak{j} = (\alpha_1, \alpha_2, \dots, \alpha_n)$$

befinden. $\mathfrak{j}$ bedeutet hier zunächst die Gesamtheit aller Zahlen

$$\lambda_1\alpha_1 + \lambda_2\alpha_2 + \dots + \lambda_n\alpha_n,$$

und es muß erst entschieden werden, ob $\alpha_1, \alpha_2, \dots, \alpha_n$ wirklich eine Basis des Ideals bilden. Endlich muß auch entschieden werden, welche von den Idealen verschieden sind.

Eine andere Methode liefert der Satz[3]):

Jedes Ideal $\mathfrak{j}$ läßt sich in der Form (α, β) schreiben, wo β eine beliebige Zahl von $\mathfrak{j}$ ist.

Wenn wir hier $\beta = N\mathfrak{j}$ wählen, so ergibt sich, daß jedes Ideal mit der Norm N sich unter den Idealen

$$(\alpha, N) \tag{4}$$

befindet, wo

$$\alpha = a_1\omega_1 + a_2\omega_2 + \ldots + a_n\omega_n$$

mit

$$0 \leqq a_i < N. \tag{5}$$

Man hat nur zu untersuchen, welche von diesen Idealen verschieden sind und die Norm N haben.

Um ein gegebenes Ideal $\mathfrak{j}$ auf die Form $(\alpha, N\mathfrak{j})$ zu bringen, bestimmt man die Norm $N\mathfrak{j} = N$ und entscheidet, welches von den Idealen (4) mit der Nebenbedingung (5) mit $\mathfrak{j}$ identisch ist.

9. Bestimmung der ganzen Zahlen mit gegebener Norm. Es sei α eine ganze Zahl und $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_r$ ein System von Fundamentaleinheiten. Dann beweist man leicht den folgenden Satz[4]):

Unter den zu α assoziierten Zahlen gibt es eine Zahl β, so daß für jedes i in der Reihe $i = 1, 2, \ldots, n$ eine Gleichung

$$\log\left|\frac{\beta^{(i)}}{\sqrt[n]{N(\beta)}}\right| = \sum_{k=1}^{r} c_k \log|\varepsilon_k^{(i)}|$$

besteht, mit reellen Koeffizienten c_k, die den Ungleichungen

$$-\frac{1}{2} < c_k \leqq \frac{1}{2} \qquad (k = 1, 2, \ldots, r)$$

genügen. (Wenn μ die Anzahl der Einheitswurzeln des Körpers bedeutet, gibt es sogar genau μ solche Zahlen β.)

Hieraus ergibt sich sofort der Satz:

Unter den ganzen Zahlen mit gegebener Norm N gibt es eine endliche Anzahl nicht-assoziierter Zahlen $\beta_1, \beta_2, \ldots, \beta_\nu$, so daß die sämtlichen ganzen Zahlen mit der Norm N durch die Zahlen

$$\varepsilon\beta_1, \varepsilon\beta_2, \ldots, \varepsilon\beta_\nu$$

gegeben sind, wenn ε die sämtlichen Einheiten des Körpers mit positiver Norm durchläuft. Diese Zahlen β können so gewählt werden, daß sie den

---

[3]) Siehe z. B. E. Landau, loc. cit., S. 150.

[4]) Siehe z. B. E. Hecke, Vorles. über die Theorie der algebraischen Zahlen (Leipzig 1923), S. 156.

folgenden Ungleichungen genügen

$$(6) \qquad \left| \log \left| \frac{\beta^{(i)}}{\sqrt[n]{N}} \right| \right| \leqq \frac{1}{2} \sum_{k=1}^{r} \left| \log | \varepsilon_k^{(i)} | \right| \qquad (1 \leqq i \leqq n).$$

Es gibt nur endlich viele ganze Zahlen β mit der Norm N, die diesen Ungleichungen genügen. Wenn wir ein System von Fundamentaleinheiten kennen, können wir sie berechnen. Gibt es keine solche Zahl, so gibt es keine ganze Zahl mit der Norm N. Unter den so gefundenen Zahlen β müssen wir entscheiden, welche miteinander assoziiert sind. Die Anzahl der Nichtassoziierten ist ν. Es ist übrigens nicht notwendig, ein System von Fundamentaleinheiten zu berechnen. Denn wegen den Relationen (1), (2) und (2′) in Nummer 5 genügt es, ein System von r unabhängigen Einheiten zu kennen.

Zwei Zahlen α und β sind dann und nur dann assoziiert, wenn $N(\alpha) = N(\beta)$ ist und wenn $\frac{\alpha}{\beta}$ eine ganze Zahl ist. Dies kann immer entschieden werden.

10. Äquivalenz zweier Ideale. Es seien gegeben die zwei Ideale $\mathfrak{a}$ und $\mathfrak{b}$, und es soll entschieden werden, ob sie äquivalent sind oder nicht. Es ist $\frac{N\mathfrak{a}}{\mathfrak{a}} = \mathfrak{c}$ ein Ideal, das wir leicht bestimmen können. Die Frage wird dann, zu entscheiden, ob das Ideal $\mathfrak{b}\,\mathfrak{c}$ ein Hauptideal ist oder nicht. Wir bestimmen die Norm N dieses Ideals und berechnen, wie in der vorigen Nummer gezeigt wurde, ein System von Zahlen $\beta_1, \beta_2, \ldots, \beta_\nu$, mit der Norm N. Gibt es keine solche Zahlen β, so ist das Ideal kein Hauptideal. Gibt es solche Zahlen β, so hat man weiter zu untersuchen, ob das Ideal $\mathfrak{b}\,\mathfrak{c}$ mit irgendeinem von den Idealen (β_i) identisch ist. Ist dies der Fall, so ist $\mathfrak{b}\,\mathfrak{c}$ ein Hauptideal, sonst nicht.

($\mathfrak{c}$ bestimmt man so: In dem von allen konjugierten Körpern gebildeten Galoisschen Körper berechnen wir das Produkt von den zu $\mathfrak{a}$ konjugierten Idealen

$$\mathfrak{a}^{(2)}\,\mathfrak{a}^{(3)} \ldots \mathfrak{a}^{(n)}.$$

Dieses Ideal liegt wieder in dem Ausgangskörper und ist gleich $\mathfrak{c}$.)

11. Aufstellung der Idealklassen. In jeder Idealklasse gibt es ein Ideal $\mathfrak{a}$, dessen Norm $\leqq |\sqrt{D}|$ ist, wenn D die Körperdiskriminante bedeutet. Wir bestimmen nun (nach Nummer 8) alle Ideale mit einer Norm $\leqq |\sqrt{D}|$ und untersuchen (nach Nummer **10**), welche von diesen äquivalent sind. Die Anzahl der nichtäquivalenten Ideale ist die Klassenzahl h. Für jede Idealklasse können wir einen Repräsentanten herauswählen.

Wie man sieht, wird es hier notwendig, um die Idealklassen aufzustellen, ein System von r unabhängigen Einheiten zu berechnen. Für die

numerische Berechnung wäre es aber bequemer, eine von den Einheiten unabhängige Methode zu haben.

Die Eigenschaften der Gruppe der Idealklassen kann man jetzt vollständig bestimmen.

12. Zerlegung der rationalen Primzahlen. Um eine gegebene rationale Primzahl in Primideale zu zerlegen, kann man sich des bekannten Satzes von Dedekind über die Zerlegung der Fundamentalgleichung, mit der Ergänzung von Hensel, bedienen. Man kann aber auch die Aufgabe auf die folgende Weise lösen: Es sei p die zu zerlegende Primzahl. Wir bestimmen nach der Nummer 8 alle Ideale mit der Norm p. Diese seien die Primideale ersten Grades

$$\mathfrak{P}_1, \mathfrak{P}_2, \ldots, \mathfrak{P}_s.$$

Wir berechnen darauf alle Ideale mit der Norm p^2, die von den Idealen $\mathfrak{P}_i \mathfrak{P}_j$ $(1 \leqq i \leqq s,\ 1 \leqq j \leqq s)$ verschieden sind. Wir erhalten dann alle Primideale zweiten Grades, die in p aufgehen,

$$\mathfrak{P}_{s+1}, \mathfrak{P}_{s+2}, \ldots, \mathfrak{P}_{s+t}.$$

Auf diese Weise können wir offenbar fortsetzen, bis wir alle Primideale, die in p aufgehen, bestimmt haben.

(Eingegangen am 1. Februar 1930.)

Über den funktionalen, insbesondere den analytischen Charakter der Lösungen elliptischer Differentialgleichungen zweiter Ordnung.

Von

Eberhard Hopf in Berlin-Dahlem.

Einleitung.

Es sei

$$\Phi\left(\frac{\partial^2 u}{\partial x_1^2}, \frac{\partial^2 u}{\partial x_1 \partial x_2}, \ldots, \frac{\partial^2 u}{\partial x_n^2}; \frac{\partial u}{\partial x_1}, \ldots, \frac{\partial u}{\partial x_n}; u; x_1, \ldots, x_n\right) = 0$$

eine partielle Differentialgleichung zweiter Ordnung in n unabhängigen reellen Veränderlichen $x_1, x_2, \ldots, x_n$. Wie allgemein bekannt, ist das Verhalten der quadratischen Form

$$\sum_{\nu \leqq \mu} \frac{\partial \Phi}{\partial \frac{\partial^2 u}{\partial x_\nu \partial x_\mu}} z_\nu z_\mu$$

von entscheidender Bedeutung für die Theorie der Gleichung $\Phi = 0$. Symmetrisch geschrieben lautet diese Form

$$\sum_{\nu,\mu} a_{\nu\mu} z_\nu z_\mu; \qquad a_{\nu\nu} = \frac{\partial \Phi}{\partial \frac{\partial^2 u}{\partial x_\nu^2}}, \qquad a_{\nu\mu} = \frac{1}{2} \frac{\partial \Phi}{\partial \frac{\partial^2 u}{\partial x_\nu \partial x_\mu}} \qquad (\nu \neq \mu)$$

(bei der Summation laufen ν, μ unabhängig voneinander). Die Gleichung $\Phi = 0$ heißt in bezug auf eine vorgegebene Lösung $u(x_1, \ldots, x_n)$ elliptisch an der Stelle $x_1 = x_1^0, \ldots, x_n = x_n^0$, wenn jene quadratische Form an dieser Stelle definit ist. Unbeschadet der Allgemeinheit kann sie dann als positiv definit angenommen werden.

Im elliptischen, und nur in diesem Falle beeinflußt der funktionale Charakter der Funktion Φ, wie man weiß, wesentlich den funktionalen Charakter der Lösungen $u(x_1, \ldots, x_n)$ von $\Phi = 0$. Diesen Zusammenhängen, welche

m Bernsteinschen Theorem[1]) vom analytischen Charakter der dreimal stetig differenzierbaren Lösungen gipfeln, ist die vorliegende Arbeit gewidmet.

Im ersten Kapitel werden Differenzierbarkeitssätze über die Lösungen von $\Phi = 0$ bewiesen. In diesem Zusammenhange sind vor allem die zuerst von Herrn Lichtenstein[2]) für $n = 2$ bewiesenen „Ergänzungssätze" zum Analytizitätstheorem zu erwähnen, die darauf abzielen, den analytischen Charakter bei besonderer Gestalt von Φ unter möglichst geringen Differenzierbarkeits-Voraussetzungen über u zu beweisen. Sätze dieser Art werden im folgenden auf Grund einer einheitlichen Methode hergeleitet, die auf der Verwendung eines meines Wissens neuen Systems von Integralformeln für die Lösungen der linearen Gleichung,

$$\Phi = \sum_{\nu,\mu} a_{\nu\mu} \frac{\partial^2 u}{\partial x_\nu \partial x_\mu} + \sum_\nu b_\nu \frac{\partial u}{\partial x_\nu} + c\,u + d,$$

beruht (die Koeffizienten $a_{\nu\mu}, b_\nu, c, d$ sind Funktionen der x_i allein). Diese Formeln erhält man durch Übergang von konstanten zu variablen Koeffizienten mit Hilfe einer zu anderem Zwecke bereits von Herrn A. Korn verwandten Zerlegung[3]). Die Ausnützung jener Formeln geschieht auf Grund leichter Ausdehnungen klassischer Überlegungen der elementaren Potentialtheorie (Sätze von O. Hölder und A. Korn). Mit Hilfe einer Schrankenbestimmung[4]) für die zweiten Ableitungen der Lösungen im linearen Falle konnte die Voraussetzung der dreimaligen stetigen Derivierbarkeit im Analytizitätssatze durch die geringere der Hölder-Stetigkeit der zweiten Ableitungen ersetzt werden[5]). Ob allgemein die bloße Stetigkeit ausreicht, ist eine zur Zeit noch offene Frage[6]).

Das zweite Kapitel ist dem Analytizitätssatze selbst gewidmet. Den älteren Beweisen der Herren S. Bernstein[7]) und M. Gevrey[8]) — es sei auch auf die Arbeit von Herrn T. Radó[9]) hingewiesen — hat neuerdings Herr

[1]) Vgl. den Encyklopädiebericht von Herrn Lichtenstein, II C 12. Betreffs des Ideenkreises um das Analytizitätstheorem vgl. S. 1320–1324.

[2]) Vgl. L. Lichtenstein, Bulletin international de l'Académie Polonaise des Sciences et des Lettres 1913, S. 915–941.

[3]) Vgl. A. Korn, Schwarz-Festschrift, Berlin 1914, S. 215–229.

[4]) Bezüglich derartiger Dinge vgl. auch Gevrey, Annales de l'Éc. Norm. **35** (1918), insbesondere S. 145–148.

[5]) Im wesentlichen mit gleicher Methode hatte ich ein entsprechendes Resultat für die Lösungen regulärer Variationsprobleme bewiesen. Vgl. Math. Zeitschr. **30** (1929), S. 404–413.

[6]) Über die Bedeutung der entsprechenden Frage bei regulären Variationsproblemen vgl. A. Haar, Abh. des Hamburger Math. Sem. 8 (1930), S. 1–27.

[7]) Vgl. loc. cit. [1]) und die neue Arbeit in Math. Zeitschr. **28** (1928), S. 330–348.

[8]) Vgl. loc. cit. [1]) und Bull. des sciences math. **50** (1926), S. 113–128.

[9]) Math. Zeitschr. **25** (1926), S. 514–589.

H. Lewy[10]) einen durchsichtigeren Beweis hinzugefügt, der auf der direkten Konstruktion der komplexen Fortsetzung der Lösung mit Hilfe der Auflösung einer hyperbolischen Anfangswertaufgabe beruht. Gegenüber diesen nur für $n = 2$ erbrachten Beweisen sei der von Herrn G. Giraud[11]) für beliebiges n geführte Beweis erwähnt. Er beruht auf eingehenden Untersuchungen über die erste Randwertaufgabe der linearen Gleichung und der analytischen Fortsetzung der Lösungen. Der Beweis des Analytizitätssatzes wird auf Grund dieser Untersuchungen durch sukzessive Approximationen unter wiederholter Auflösung erster Randwertaufgaben erbracht. Im folgenden wird der Bernsteinsche Satz für beliebige n in anderer Weise bewiesen. Die Gleichung $\Phi = 0$ wird zunächst in ein nichtlineares Integralgleichungssystem übergeführt. Mit Hilfe eines wichtigen Gedankens von Herrn E. E. Levi[12]), den in anderer Richtung auch Herr Giraud ausbeutete, lassen sich die Integrale dieses Gleichungssystems auch für komplexe Werte der Variablen definieren. Das so erhaltene komplexe Gleichungssystem, das jedenfalls von der analytischen Fortsetzung der Lösung von $\Phi = 0$ befriedigt werden muß, wird gelöst, und die Regularität der Lösung nachgewiesen. Ihrer Natur nach kann die Methode auf eine ausgedehnte Klasse nichtlinearer Integral- und Integro-Differentialgleichungen angewandt werden. Es sei noch bemerkt, daß das II. Kapitel nur die Lektüre von § 1 voraussetzt[13]).

Inhaltsübersicht.

Seite

I. Kapitel. Funktionale Eigenschaften der Lösungen elliptischer Differentialgleichungen zweiter Ordnung.

§ 1 Greensche Formeln. Die Funktion $\Gamma(a; Z)$ 197
§ 2. Hilfssätze über „Potentiale“ n-dimensionaler Belegungen . 202
§ 3. Differenzierbarkeitssätze über die Lösungen elliptischer Differentialgleichungen 206
§ 4. Verschärfung des Satzes 4 211

[10]) Math. Annalen **101** (1929), S. 609–619.

[11]) Annales de l'Éc. Norm. **43** (1926), S. 1–128.

[12]) Palermo Rendiconti **24** (1907), p. 275–317, insbesondere p. 297–307.

[13]) Sämtliche hier entwickelten Methoden sind durchaus auf elliptische Gleichungen höherer Ordnung, sowie auf Systeme erster Ordnung anwendbar. Der einzige springende Punkt bei dieser Ausdehnung besteht in der Konstruktion einer Grundlösung einer beliebigen linearen elliptischen Gleichung mit konstanten Koeffizienten, bei welcher nur die Ableitungen einer festen Ordnung vorkommen. Im Falle zweiter Ordnung ist dies gerade die Funktion $\Gamma(a; Z)$ von § 1. Im Falle $n = 2$ und beliebiger (natürlich gerader) Ordnung vgl. loc. cit.[12]), S. 280–284. Im allgemeinen Falle sind die wichtigen Arbeiten von Herrn Herglotz heranzuziehen. Vgl.: Abh. des Hamburger Math. Sem. **6** (1928), S. 189–197; Berichte der Sächsischen Akademie **78** (1926), S. 93–126; **78** (1926), S. 287–318; **80** (1928), S. 69–114.

Fortsetzung der Inhaltsübersicht.

Seite

II. Kapitel. Beweis des analytischen Charakters der Lösungen von $\Phi = 0$.
§ 5. Ein System nichtlinearer Integralgleichungen für die Lösungen von $\Phi = 0$ 216
§ 6. Vorbereitungen zum Übergang ins Komplexe 220
§ 7. Definition und Eigenschaften des Integrales in (6.6) für komplexe Punkte X 223
§ 8. Das komplexe Integralgleichungssystem. Existenz und Regularität des Lösungssystems 228

I. Kapitel.

Funktionale Eigenschaften der Lösungen elliptischer Differentialgleichungen zweiter Ordnung.

§ 1.

Greensche Formeln. Die Funktion $\Gamma(a; Z)$.

Mit X, Y, Z usw. seien im folgenden Punkte des n-dimensionalen kartesischen Raumes mit den Koordinaten x_ν, y_ν, z_ν usw., $\nu = 1, 2, \ldots, n$ bezeichnet. Es ist klar, was unter $aX + bY$ zu verstehen ist (a, b Konstanten). Des weiteren sei das Symbol $|X| = \sqrt{\sum_\nu x_n^2}$ benützt. Ist $\mathfrak{B}$ ein Bereich[14]) in diesem Raume und $f(X)$ eine in $\mathfrak{B}$ definierte Funktion, so sei kurz

$$\int\limits_{\mathfrak{B}}^{(n)} \ldots \int f(X)\, dx_1\, dx_2 \ldots dx_n = \int\limits_{\mathfrak{B}}^{(n)} f(X)\, dX$$

geschrieben. Ist $\mathfrak{B}$ ein Hyperkugel-Bereich $|X| = r \leqq R$, so ergibt sich, wenn $f(X) = \varphi(r)$ nur von r abhängt, leicht

$$\int\limits_{r<R}^{(n)} \varphi(r)\, dX = n\, e_n \int\limits_0^R r^{n-1} \varphi(r)\, dr, \tag{1.1}$$

unter e_n im folgenden stets den Inhalt von $|X| \leqq 1$ verstanden. Speziell ist also

$$\int\limits_{R_1<r<R_2}^{(n)} \frac{dX}{r^\lambda} = n\, e_n \frac{R_2^{n-\lambda} - R_1^{n-\lambda}}{n - \lambda}. \tag{1.2}$$

Demnach konvergiert das Integral über eine bis auf eine Stelle $X = X_0$ in $\mathfrak{B}$ stetige Funktion $f(X)$ absolut, wenn $f(X)$ an dieser Stelle von kleinerer als n-ter Ordnung unendlich wird.

---

[14]) Im allgemeinen sei an üblichen Definitionen festgehalten: Gebiet = zusammenhängende offene Punktmenge, Bereich = Gebiet + Rand des Gebietes.

Der Bereich $\mathfrak{B}$ sei konvex, $\mathfrak{R}$ sei sein Rand. Bedeutet $\mathfrak{B}_i$ im Moment die Projektion von $\mathfrak{R}$ auf die Hyperebene $x_i = 0$, so lassen sich die Punkte von $\mathfrak{B}$ vollständig durch zwei Ungleichungen

$$g_i(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \leqq x_i \leqq G_i(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$$

charakterisieren, unter g_i, G_i zwei in $\mathfrak{B}_i$ stetige Funktion verstanden. Ist $f(X)$ auf $\mathfrak{R}$ definiert, so sei

$$\int_{\mathfrak{B}_i}^{(n-1)} \{f_{x_i=G_i} - f_{x_i=g_i}\}\, dx_1 \ldots dx_{i-1}\, dx_{i+1} \ldots dx_n = \int_{\mathfrak{R}}^{(n-1)} f(X)\, d_i X$$

für die Randintegrale geschrieben. Die Verwandlungsformel von Bereichintegralen in Randintegrale lautet dann

$$\int_{\mathfrak{B}}^{(n)} \frac{\partial f(X)}{\partial x_i}\, dX = \int_{\mathfrak{R}}^{(n-1)} f(X)\, d_i X. \tag{1.3}$$

(1.3) gilt natürlich auch, wenn $\mathfrak{B}$ ein von zwei konvexen Hyperflächen begrenzter Bereich ist; dann ist nur $\int_{\mathfrak{R}} = \int_{\mathfrak{R}_a} - \int_{\mathfrak{R}_i}$ zu setzen, wo $\mathfrak{R}_a$ den äußeren, $\mathfrak{R}_i$ den inneren Rand von $\mathfrak{B}$ zu bedeuten hat. Ist $\mathfrak{B}$ wieder konvex, R der Radius der kleinsten $\mathfrak{B}$ umschließenden Hyperkugel, so gilt offenbar

$$\left| \int_{\mathfrak{R}}^{(n-1)} f(X)\, d_i X \right| \leqq 2 e_{n-1} R^{n-1} \operatorname*{Max}_{X \text{ auf } \mathfrak{R}} |f(X)|. \tag{1.4}$$

(1.3) gilt noch, wenn die Stetigkeit von f und $\frac{\partial f}{\partial x_i}$ in einem inneren Punkte X_0 von $\mathfrak{B}$ unterbrochen ist, in welchem f von kleinerer als $(n-1)$-ter, $\frac{\partial f}{\partial x_i}$ von kleinerer als n-ter Ordnung unendlich werden. In bekannter Weise sieht man dies, wenn man aus $\mathfrak{B}$ das Innere einer kleinen Hyperkugel $|X - X_0| = r$ entfernt und nach Anwendung von (1.3) und (1.4) $r \to 0$ streben läßt.

Es sei nun

$$L(u) = L_X(u) \equiv \sum_{\nu,\mu} a_{\nu\mu}(X) \frac{\partial^2 u}{\partial x_\nu \partial x_\mu}; \qquad a_{\nu\mu} \equiv a_{\mu\nu}$$

ein linearer partieller Differentialausdruck zweiter Ordnung,

$$M(u) = M_X(u) \equiv \sum_{\nu,\mu} \frac{\partial^2 (a_{\nu\mu} u)}{\partial x_\nu \partial x_\mu}$$

der adjungierte Ausdruck. Integriert man die bekannte Identität

$$v\, L(u) - u\, M(v) \equiv \sum_{\nu} \frac{\partial N_\nu}{\partial x_\nu},$$

$$N_\nu = N_\nu\left(u, v, \frac{\partial u}{\partial X}, \frac{\partial v}{\partial X}\right) = \sum_{\lambda} \left[a_{\lambda\nu} \left(v \frac{\partial u}{\partial x_\lambda} - u \frac{\partial v}{\partial x_\lambda} \right) - u v \frac{\partial a_{\lambda\nu}}{\partial x_\lambda} \right]$$

über einen konvexen Bereich $\mathfrak{B}$ — in einem $\mathfrak{B}$ enthaltenden Gebiete seien $a_{\nu\mu}(X)$, $u(X)$, $v(X)$ zweimal stetig differenzierbar —, so folgt die Greensche Formel

$$(1.5)\quad \int_{\mathfrak{B}}^{(n)} \{v(Y)\, L_Y(u) - u(Y)\, M_Y(v)\}\, dY = \sum_{\nu} \int_{\mathfrak{R}}^{(n-1)} N_\nu\Big(u, v, \frac{\partial u}{\partial Y}, \frac{\partial v}{\partial Y}\Big)\, d_\nu Y.$$

Sie bleibt nach dem Obigen gültig, wenn die zweiten Ableitungen $\frac{\partial^2 v}{\partial Y^2}$ in $\mathfrak{B}$ stetig sind bis auf einen inneren Punkt, in welchem sie von kleinerer als der n-ten Ordnung unendlich werden.

Nunmehr sei $L(u)$ im Bereiche $\mathfrak{B}$ vom *elliptischen* Typus, d. h. die charakteristische quadratische Form $\sum_{\nu,\mu} a_{\nu\mu}(X) z_\nu z_\mu$ sei in jedem Punkte X von $\mathfrak{B}$ positiv definit. Die Diskriminante $\Delta = |a_{ik}(X)|$ sei gleich Eins für alle X vorausgesetzt; ist dies nicht erfüllt, so muß man vorher immer erst $L(u)$ durch $\Delta^{\frac{1}{n}}$ dividieren, um auf diesen Fall zurückzukommen[15]). Dann ist bekanntlich auch die reziproke quadratische Form

$$\Omega(a; Z) = \sum_{\nu,\mu} A_{\nu\mu} z_\nu z_\mu$$

($A_{\nu\mu}$ = algebraisches Komplement von $a_{\nu\mu}$) in $\mathfrak{B}$ positiv definit, und ihre Diskriminante in $\mathfrak{B}$ ebenfalls gleich Eins. Die Schreibweise $\Omega(a; Z)$ soll die unmittelbare Abhängigkeit von den $a = a_{\nu\mu}$ andeuten. In unserem Falle wäre genauer $\Omega(a(X); Z)$ zu schreiben. Aus Stetigkeitsgründen gibt es ein $\omega > 0$, so daß

$$(1.6)\quad \Omega(a(X); Z) \geqq \omega |Z|^2$$

für alle X in $\mathfrak{B}$ gilt. Wir betrachten die Funktion

$$\Gamma(a; Z) = \frac{\Omega(a; Z)^{-\frac{n-2}{2}}}{n(n-2)e_n}.$$[16])

Sie hat die folgenden Eigenschaften. Aus den Formeln

$$(1.7)\quad 2n e_n \frac{\partial \Gamma}{\partial z_\nu} = -\Omega^{-\frac{n}{2}} \frac{\partial \Omega}{\partial z_\nu}, \qquad 2n e_n \frac{\partial^2 \Gamma}{\partial z_\nu \partial z_\mu} = \Omega^{-\frac{n+2}{2}} \Big(\frac{n}{2} \frac{\partial \Omega}{\partial z_\nu} \frac{\partial \Omega}{\partial z_\mu} - \Omega \frac{\partial^2 \Omega}{\partial z_\nu \partial z_\mu}\Big)$$

[15]) Bei allen Anwendungen denke man sich im folgenden vorher diese Division vorgenommen. Sie ändert nirgends etwas an den jeweiligen Voraussetzungen.

[16]) Für $n = 2$ hat man den bekannten Logarithmusausdruck. Die im folgenden gemachte Voraussetzung $n \geqq 3$ bedeutet keinen wirklichen Ausschluß des Falles $n = 2$; denn alle Sätze über die Lösungen elliptischer Differentialgleichungen in n unabhängigen Variablen enthalten die entsprechenden Sätze für weniger als n Variable, wie man sehr leicht einsieht.

erkennt man mit Rücksicht auf die bekannten Eigenschaften der reziproken Form[17]) leicht, daß Γ der Gleichung

$$\sum_{\nu,\mu} a_{\nu\mu} \frac{\partial^2 \Gamma(a;Z)}{\partial z_\nu \partial z_\mu} = 0 \tag{1.8}$$

genügt. Ferner werden nach (1.7) für $Z = 0$ Γ von $(n-2)$-ter, die ersten Ableitungen $\frac{\partial \Gamma}{\partial Z}$ von $(n-1)$-ter und die zweiten Ableitungen $\frac{\partial^2 \Gamma}{\partial Z^2}$ von n-ter Ordnung unendlich.

Unter Einführung von

$$G(X, Y) = \Gamma(a(X); Y - X) \tag{1.9}$$

zeigen wir kurz in bekannter Weise, daß jede in einem $\mathfrak{B}$ enthaltenden Gebiete zweimal stetig differenzierbare Funktion $u(X)$ für jeden inneren Punkt X von $\mathfrak{B}$ der Integralformel

$$u(X) = \int_{\mathfrak{B}}^{(n)} \{u(Y) M_Y(G) - G L_Y(u)\} d Y + \sum_\nu \int_{\mathfrak{R}}^{(n-1)} N_\nu \left(u, G, \frac{\partial u}{\partial Y}, \frac{\partial G}{\partial Y}\right) d_\nu Y \tag{1.10}$$

genügt.

Man bemerkt zunächst, daß $M_Y(G)$ für $Y = X$ nur von $(n-1)$-ter Ordnung unendlich wird. Dies ist sofort ersichtlich, wenn man den wegen (1.8) und (1.9) verschwindenden Ausdruck

$$\sum_{\nu,\mu} a_{\nu\mu}(X) \frac{\partial^2 G}{\partial y_\nu \partial y_\mu}$$

abzieht. Zum Beweise von (1.10) entferne man aus $\mathfrak{B}$ das Innere eines kleinen Hyperellipsoides

$$\Omega(a(X);\ Y - X) = r^2$$

mit dem festgehaltenen Punkte X als Mittelpunkt und wende man auf den Restbereich $\mathfrak{B}'$ die Greensche Formel mit $v(Y) = G(X, Y)$ an,

$$\int_{\mathfrak{B}'}^{(n)} \{G L_Y(u) - u M_Y(G)\} d Y = \sum_\nu \int_{\mathfrak{R}}^{(n-1)} N_\nu \left(u, G, \frac{\partial u}{\partial Y}, \frac{\partial G}{\partial Y}\right) - J_1 - J_2,$$

$$J_1 = -\sum_{\nu,\mu} \int_{\Omega = r^2}^{(n-1)} u \frac{\partial G}{\partial y_\mu} a_{\nu\mu} d_\nu Y, \qquad J_2 = \sum_{\nu,\mu} \int_{\Omega = r^2}^{(n-1)} \left(G \frac{\partial u}{\partial y_\mu} a_{\nu\mu} - u G \frac{\partial a_{\nu\mu}}{\partial y_\mu}\right) d_\nu Y.$$

Man sieht wegen $G = O(r^{2-n})$ auf $\Omega = r^2$ zunächst, daß für $r \to 0$ $J_2 \to 0$ strebt. Wir zeigen, daß für $r \to 0$ $J_1 \to u(X)$ konvergiert. Auf dem Hyperellipsoid ist nun wegen (1.7)

$$-\frac{\partial G}{\partial y_\mu} = \frac{r^{-n}}{2n e_\kappa} \frac{\partial \Omega}{\partial y_\mu},$$

[17]) Vgl. G. Kowalewski, Determinantentheorie, Leipzig 1909, § 103.

also, wenn man in Bereichintegrale zurückverwandelt,

$$2n e_n J_1 = r^{-n} \sum_{\nu,\mu} \int_{\Omega=r^2}^{(n-1)} a_{\nu\mu} u \frac{\partial \Omega}{\partial y_\mu} d_\nu Y = r^{-n} \int_{\Omega<r^2}^{(n)} \left\{ \sum_{\nu,\mu} \frac{\partial^2 (u a_{\nu\mu} \Omega)}{\partial y_\nu \partial y_\mu} \right\} dY.$$

Der Integrand im rechten Integral hat für $Y = X$ den Wert

$$2u(X) \sum_{\nu,\mu} a_{\nu\mu} A_{\nu\mu} = 2n u(X),$$

somit folgt

$$J_1 = \frac{u(X)}{e_n r^n} \int_{\Omega<r^2}^{(n)} dY + O(r) = \frac{u(X)}{e_n} \int_{\Omega<1}^{(n)} dY + O(r).$$

Da die Diskriminante der Form Ω gleich Eins ist, muß die Determinante der linearen, das Gebiet $\Omega < 1$ in das Innere der Einheits-Hyperkugel überführenden Transformation gleich ± 1 sein. Somit ist der Inhalt des Gebietes $\Omega < 1$ gleich e_n, und man hat in der Tat $J_1 \to u(X)$ für $r \to 0$.

Die Integralformel (1.10) bleibt gültig, wenn G durch eine Funktion $G' = G + v$ ersetzt, sobald nur die zweiten Ableitungen $\frac{\partial^2 v}{\partial Y^2}$ für $Y = X$ von höchstens $(n-1)$-ter Ordnung unendlich werden; denn für u und v (bei festgehaltenem X) gilt, wie erwähnt, die Greensche Formel (1.5). Im zweiten Kapitel wird z. B. die diesen Bedingungen gewiß genügende Funktion $H(X, Y) = G(Y, X)$ verwendet.

Mit Rücksicht auf spätere Anwendungen werde die Funktion $\Gamma(a; Z)$ noch genauer untersucht. Von (1.7) ausgehend erkennt man leicht, daß die partiellen Ableitungen $\frac{\partial^i}{\partial Z^i}$ die Form

$$\frac{\partial^i \Gamma(a; Z)}{\partial Z^i} = \frac{P_i(a; Z)}{\Omega(a; Z)^{\frac{n+2i-2}{2}}} \tag{1.11}$$

besitzen, unter P_i ein homogenes Polynom i-ten Grades in den z_ν verstanden, dessen Koeffizienten selbst Polynome in den $a = a_{\nu\mu}$ mit konstanten Koeffizienten sind. (1.11) kann man auch in der Gestalt

$$Z|^{n+i-2} \frac{\partial^i \Gamma(a; Z)}{\partial Z^i} = \frac{P(a; Z|Z|^{-1})}{\Omega(a; Z|Z|^{-1})^k} \tag{1.12}$$

schreiben, wo $Z|Z|^{-1}$ das Wertesystem mit den Komponenten $z_\nu|Z|^{-1}$ (ihre Quadratsumme ist gleich Eins) bedeutet und k eine positive Zahl ist; der Zähler P ist jedenfalls ein Polynom aller seiner Argumente mit konstanten Koeffizienten. Differenziert man überdies in (1.12) nach den $a_{\nu\mu}$ — dies sei kurz mit $\frac{\partial}{\partial a}$ angedeutet —, so erkennt man sofort, daß dabei

die Form des Ausdrucks (1.12) erhalten bleibt,

$$(1.13) \qquad |Z|^{n+i-2}\cdot\frac{\partial^{i+k}\Gamma(a;Z)}{\partial Z^i\partial a^k}=\frac{P'(a;Z|Z|^{-1})}{\Omega(a;Z|Z|^{-1})^{k'}} \qquad (k'>0).$$

Man erkennt jedenfalls aus dieser Darstellung, daß sich irgendwelche Stetigkeitseigenschaften, z. B. Hölder-Stetigkeit, der Koeffizientenfunktionen $a_{\nu\mu}=a_{\nu\mu}(X)$ im obigen Bereiche $\mathfrak{B}$ automatisch auf die linke Seite in (1.13) übertragen, und zwar gleichmäßig bezüglich Z, weil der Nenner nach (1.6) $\geqq\omega^{k'}$ ist.

Mit Rücksicht auf eine Anwendung im zweiten Kapitel sei noch die ebenfalls leicht von (1.11) aus zu konstatierende Bemerkung gestattet, daß die gemischten Ableitungen genauer von der Form

$$(1.14) \qquad \frac{\partial^{i+k}\Gamma(a;Z)}{\partial Z^i\partial a^k}=\frac{P_{i+2k}(a;Z)}{\Omega(a;Z)^{\frac{n+2i+2k-2}{2}}}$$

sind, wo P_{i+2k} homogen vom $(i+2k)$-ten Grade in den z_ν ist.

Von der obigen Funktion $G(X,Y)$ von (1.9) gilt in leicht verständlicher symbolischer Schreibweise. $Z=Y-X$,

$$\frac{\partial^i G}{\partial Y^i}=\frac{\partial^i\Gamma}{\partial Z^i},\quad \frac{\partial^{i+1}G}{\partial Y^i\partial X}=\sum\frac{\partial^{i+1}\Gamma}{\partial Z^i\partial a}\frac{\partial a}{\partial X}+\frac{\partial^{i+1}\Gamma}{\partial Z^{i+1}}$$

Es gelten also wegen (1.13) Ungleichungen der Form

$$(1.15) \qquad \left|\frac{\partial^i G}{\partial Y^i}\right|<\frac{\text{konst.}}{|Y-X|^{n+i-2}},\qquad \left|\frac{\partial^{i+1}G}{\partial Y^i\partial X}\right|<\frac{\text{konst.}}{|Y-X|^{n+i-1}}$$

für alle Punkte X, Y aus $\mathfrak{B}$. Die linke Ungleichung setzt nur die Stetigkeit der $a_{\nu\mu}(X)$ in $\mathfrak{B}$ voraus.

§ 2.

Hilfssätze über „Potentiale“ n-dimensionaler Belegungen.

Für die weiteren Untersuchungen sind einige Zwischenbetrachtungen über Funktionen der Form

$$f(X)=\int_{\mathfrak{B}}^{(n)}\varrho(Y)\,H(X,Y)\,dY$$

erforderlich. $H(X,Y)$ soll dabei eine mit Einschluß der jeweils vorkommenden partiellen Ableitungen in $\mathfrak{B}$ stetige Funktion von X und Y sein; nur für $X=Y$ darf H gewisse Singularitäten besitzen. Die „Belegung“ $\varrho(X)$ soll gewisse Stetigkeitseigenschaften besitzen, etwa einer Hölder-Bedingung genügen. Die Definition ist bekanntlich die folgende. $\varrho(X)$ ist in einem Bereiche $\mathfrak{B}$ Hölder-stetig mit dem Exponenten $\alpha>0$, wenn $|\varrho(X'')-\varrho(X')|\leqq$ konst. $|X''-X'|^\alpha$ für irgend zwei Punkte X'', X' aus $\mathfrak{B}$ gilt. ϱ heißt im Innern eines Bereiches $\mathfrak{B}$, oder in einem Gebiete Hölder-

stetig, wenn ϱ im eben erklärten Sinne in jedem mit seinen Randpunkten ganz im Innern gelegenen Teilbereiche Hölder-stetig ist.

Hilfssatz 1. Es sei für alle Punkte X und Y aus $\mathfrak{B}$

$$|H| < \frac{\text{konst.}}{|Y-X|^{n-1}}, \qquad \left|\frac{\partial H}{\partial X}\right| < \frac{\text{konst.}}{|Y-X|^{n}}.$$

Ist $\varrho(X)$ in $\mathfrak{B}$ Hölder-stetig, so ist die Funktion

$$\varphi(X) = \int_{\mathfrak{B}}^{(n)} \{\varrho(Y) - \varrho(X_0)\} H(X, Y)\, dY$$

an der Stelle $X = X_0$ (X_0 innerhalb $\mathfrak{B}$) einmal derivierbar, und es gilt an dieser Stelle

$$\frac{\partial \varphi}{\partial X} = \int_{\mathfrak{B}}^{(n)} \{\varrho(Y) - \varrho(X_0)\} \frac{\partial H}{\partial X}\, dY.$$

Den auf geläufigen potentialtheoretischen Schlußweisen beruhenden Beweis können wir uns ersparen. Eine unmittelbare Folge aus dem Hilfssatz 1 ist der

Hilfssatz 2. Sind die Voraussetzungen des Hilfssatzes 1 erfüllt, und ist die Funktion

$$\psi(X) = \int_{\mathfrak{B}}^{(n)} H(X, Y)\, dY$$

im Innern von $\mathfrak{B}$ einmal stetig differenzierbar, so besitzt die Funktion

$$f(X) = \int_{\mathfrak{B}}^{(n)} \varrho(Y) H(X, Y)\, dY$$

die gleiche Eigenschaft, und es gilt

$$\frac{\partial f(X)}{\partial X} = \varrho(X) \frac{\partial \psi(X)}{\partial X} + \int_{\mathfrak{B}}^{(n)} \{\varrho(Y) - \varrho(X)\} \frac{\partial H(X, Y)}{\partial X}\, dY.$$

Man hat nur in irgendeinem inneren Punkte X_0 von $\mathfrak{B}$ den Hilfssatz 1 auf die Funktion

$$f(X) = \varrho(X_0)\, \psi(X) + \int_{\mathfrak{B}}^{(n)} \{\varrho(Y) - \varrho(X_0)\} H(X, Y)\, dY$$

anzuwenden.

Hilfssatz 3. $K(X, Z)$ sei eine mit ihren partiellen Ableitungen $\frac{\partial K}{\partial Z}$, $\frac{\partial^2 K}{\partial Z^2}$ für alle X eines konvexen Bereichs $\mathfrak{B}$ und alle $Z \neq 0$ stetige Funktion von X, Z. Es gelte

$$|K| < \frac{\text{konst.}}{|Z|^{n-1}}, \qquad \left|\frac{\partial K}{\partial Z}\right| < \frac{\text{konst.}}{|Z|^{n}}, \qquad \left|\frac{\partial^2 K}{\partial Z^2}\right| < \frac{\text{konst.}}{|Z|^{n+1}}$$

für alle X aus $\mathfrak{B}$ und alle Z, und die Funktionen $|Z|^n \frac{\partial K}{\partial Z}$ seien in $\mathfrak{B}$

Hölder-stetige Funktionen von X mit dem Exponenten α $(0 < \alpha < 1)$, und zwar gleichmäßig bezüglich Z. Ist $\varrho(X)$ ebenfalls in $\mathfrak{B}$ Hölder-stetig mit dem Exponenten α, so ist auch

$$F(X) = \int_{\mathfrak{B}}^{(n)} \{\varrho(Y) - \varrho(X)\}\, \sigma(Y) \frac{\partial K(X, Y-X)}{\partial Y}\, dY$$

in $\mathfrak{B}$ Hölder-stetig, im allgemeinen jedoch mit einem Exponenten $< \alpha$.

Korollar. Ist außerdem noch $\sigma(X)$ in $\mathfrak{B}$ Hölder-stetig — auf den Exponenten kommt es dabei nicht an —, so ist $F(X)$ im Innern von $\mathfrak{B}$ Hölder-stetig mit dem Exponenten α.

Beweis. Man setze

$$F(X, Z) = \int_{\mathfrak{B}}^{(n)} \{\varrho(Y) - \varrho(Z)\}\, \sigma(Y) \frac{\partial K(X, Y-Z)}{\partial Y}\, dY.$$

Wegen $F(X) = F(X, X)$ genügt der Nachweis, daß $F(X, Z)$ als Funktion jedes Argumentpunktes (X, Z in $\mathfrak{B}$) Hölder-stetig ist, und zwar gleichmäßig bezüglich des anderen Argumentpunktes. Zunächst folgt aus den Voraussetzungen leicht

$$\begin{aligned} |F(X'', Z) - F(X', Z)| &\leqq \text{konst.} |X'' - X'|^{\alpha} \cdot \int_{\mathfrak{B}}^{(n)} \frac{dY}{|Y - Z|^{n-\alpha}} \\ &\leqq \text{konst.} |X'' - X'|^{\alpha} \end{aligned}$$

für alle X', X'', Z aus $\mathfrak{B}$, so daß $F(X, Z)$ nur noch als Funktion von Z zu untersuchen ist. Zu diesem Zweck sei für irgend zwei Punkte Z', Z'' von $\mathfrak{B}$

$$r = 2|Z'' - Z'| \tag{2.1}$$

gesetzt und unter K_r der Durchschnitt von $\mathfrak{B}$ mit dem Hyperkugelinnern $|Y - Z'| < r$ verstanden[18]). Setzt man zur Abkürzung noch

$$H(X, Y; Z) = \{\varrho(Y) - \varrho(Z)\}\, \sigma(Y) \frac{\partial K(X, Y-Z)}{\partial Y}, \tag{2.2}$$

so ist

$$F(X, Z'') - F(X, Z') = \{\varrho(Z'') - \varrho(Z')\}\, J + J_1 + J_2,$$

$$J = \int_{\mathfrak{B}-K_r}^{(n)} \sigma(Y) \frac{\partial K(X, Y-Z')}{\partial Y}\, dY, \quad J_1 = \int_{K_r}^{(n)} \{H(X, Y; Z'') - H(X, Y; Z')\}\, dY, \tag{2.3}$$

$$J_2 = \int_{\mathfrak{B}-K_r}^{(n)} \{\varrho(Y) - \varrho(Z'')\}\, \sigma(Y) \left\{ \frac{\partial K(X, Y-Z'')}{\partial Y} - \frac{\partial K(X, Y-Z')}{\partial Y} \right\} dY.$$

[18]) $\mathfrak{B}_r$ ist niemals leer, wenn man etwa r der Bedingung $r < D$ unterwirft, wo D der Durchmesser des Bereiches $\mathfrak{B}$ ist.

J_1 werde zuerst betrachtet. Wegen (2.2) ist nach den Voraussetzungen

$$|H(X, Y; Z)| < \frac{\text{konst.}}{|Y-Z|^{n-\alpha}}$$

für alle X, Y, Z aus $\mathfrak{B}$. Also hat man

$$|J_1| < \text{konst.} \int_{K_r}^{(n)} \frac{dY}{|Y-Z'|^{n-\alpha}} + \text{konst.} \int_{K_r}^{(n)} \frac{dY}{|Y-Z''|^{n-\alpha}}.$$

Da Z', Z'' beide in K_r liegen, vergrößert man die rechte Seite, wenn man das erste Integral über das Gebiet $|Y-Z'| < 2r$, das zweite über das Gebiet $|Y-Z''| < 2r$ erstreckt,

$$|J_1| < \text{konst.} \int_{\varrho<2r}^{(n)} \frac{dY}{\varrho^{n-\alpha}} = \text{konst.}\, r^{\alpha}. \tag{2.4}$$

Mit $\mathfrak{B}$ ist auch K_r ein konvexer Bereich; da Z', Z'' in K_r liegen, ist nach dem Mittelwertsatze

$$\Delta = \frac{\partial K(X, Y-Z'')}{\partial Y} - \frac{\partial K(X, Y-Z')}{\partial Y} = \sum_{\nu} (z_\nu'' - z_\nu') \frac{\partial^2 K(X, Y-Z)}{\partial Y \partial z_\nu}\Big|_{Z=Z^*}, \tag{2.5}$$

wo Z^* auf der Verbindungslinie zwischen Z' und Z'', also in K_r liegt. Nach den Voraussetzungen ist daher

$$|\Delta| \leqq \text{konst.} \frac{|Z''-Z'|}{|Y-Z^*|^{n+1}} = \text{konst.}\, r\, |Y-Z^*|^{-n-1}.$$

Liegt nun Y in $\mathfrak{B} - K_r$, so ist $|Y-Z'| > 2|Z''-Z'|$ und demnach $|Y-Z^*| \geqq |Y-Z'| - |Z'-Z^*| > |Y-Z'| - |Z''-Z'| > \frac{1}{2}|Y-Z'|$, also

$$|\Delta| < \text{konst.}\, r\, |Y-Z'|^{-n-1}, \quad Y \text{ in } \mathfrak{B} - K_r. \tag{2.6}$$

Des weiteren ist $|Y-Z''| < |Y-Z'| + |Z''-Z'| < \frac{3}{2}|Y-Z'|$, also

$$|\varrho(Y) - \varrho(Z'')| < \text{konst.}\, |Y-Z'|^{\alpha}, \quad Y \text{ in } \mathfrak{B} - K_r.$$

Hieraus, sowie aus (2.5) und (2.6) folgt, wenn D den Durchmesser des Bereiches $\mathfrak{B}$ bedeutet, für $r < D$

$$|J_2| < \text{konst.}\, r \int_{\mathfrak{B}-K_r}^{(n)} \frac{dY}{|X-Y'|^{n+1-\alpha}} < \text{konst.}\, r \int_{r<\varrho<D}^{(n)} \frac{dY}{\varrho^{n+1-\alpha}} = \text{konst.}\, r \frac{r^{\alpha-1} - D^{\alpha-1}}{1-\alpha},$$

also [19])

$$|J_2| < \text{konst.}\, r^{\alpha} \tag{2.7}$$

für beliebige X, Z', Z'' aus $\mathfrak{B}$. Was endlich J betrifft, so hat man nach

[19]) Hierbei wird die Voraussetzung $\alpha < 1$ benützt.

Voraussetzung für $r < D$

$$|J| < \text{konst.} \int\limits_{\mathfrak{B}-K_r}^{(n)} \frac{dY}{|Y-Z'|^n} < \text{konst.} \int\limits_{r<\varrho<D}^{(n)} \frac{dY}{|Y-Z'|^n} = \text{konst.} \log \frac{D}{r}.$$

Berücksichtigt man dieses, sowie (2.4) und (2.7) in (2.3), so folgert man, daß ($r = 2|Z''-Z'|$) $F(X, Z)$ als Funktion von Z für beliebig vorgegebenes β mit $\beta < \alpha$ in $\mathfrak{B}$ Hölder-stetig mit dem Exponenten β ist.

Beweis des Korollars. Nach dem Obigen genügt es zu zeigen, daß $|J|$ unter einer festen Schranke gelegen ist, wenn die Punkte Z', Z'' in einem ganz innerhalb von $\mathfrak{B}$ liegenden Teilbereiche $\mathfrak{B}'$ variieren, und wenn r der Bedingung

$$r = 2|Z''-Z'| < d \tag{2.8}$$

genügt, unter d die Entfernung des Randes von $\mathfrak{B}'$ vom Rande $\mathfrak{R}$ von $\mathfrak{B}$ verstanden. Nun ist

$$J = \int\limits_{\mathfrak{B}-K_r}^{(n)} \{\sigma(Y) - \sigma(Z')\} \frac{\partial K(X, Y-Z')}{\partial Y} dY + \sigma(Z') \int\limits_{\mathfrak{B}-K_r}^{(n)} \frac{\partial K(X, Y-Z')}{\partial Y} dY.$$

Da $\sigma(Y)$ in $\mathfrak{B}$ Hölder-stetig sein sollte, ist das erste Integral gewiß für alle X, Z' aus $\mathfrak{B}$ absolut unter einer festen Schranke gelegen. Für das zweite Integral kann man auch, da wegen (2.8) der Y-Bereich $|Y-Z'| \leqq r$ ganz zu $\mathfrak{B}$ gehört, $\left(\frac{\partial}{\partial Y} = \frac{\partial}{\partial y_i}\right)$

$$\int\limits_{\mathfrak{R}}^{(n-1)} K(X, Y-Z')\, d_i Y - \int\limits_{|Y-Z'|=r}^{(n-1)} K(X, Y-Z')\, d_i Y$$

schreiben. Da Z' in $\mathfrak{B}'$ liegt, ist wegen $|Y-Z'| \geqq d$ das erste Integral gewiß beschränkt. Das zweite ist ebenfalls beschränkt, wie man mit Rücksicht auf die Voraussetzung über K mit Hilfe von (1.4) sofort erkennt. Daraus folgt die Behauptung über $|J|$. $F(X, Z)$ ist also als Funktion von Z in $\mathfrak{B}'$ Hölder-stetig mit dem Exponenten α, und zwar gleichmäßig bezüglich X.

§ 3.

Differenzierbarkeitssätze über die Lösungen elliptischer Differentialgleichungen.

Wir betrachten die lineare, in einem bestimmten Gebiete $\mathfrak{G}$ elliptische Differentialgleichung zweiter Ordnung

$$L(u) = \sum_{\nu,\mu} a_{\nu\mu}(X) \frac{\partial^2 u}{\partial x_\nu \partial x_\mu} = f(x).$$

Von den Funktionen $a_{\nu\mu}(X)$ und $f(X)$ sei jetzt nur noch die Hölder-Stetigkeit im Gebiete $\mathfrak{G}$ vorausgesetzt. Unter so geringen Annahmen wird natürlich die gewöhnliche Integralformel (1.10) hinfällig. Wir behaupten nun, daß die zweiten Ableitungen jeder in G zweimal stetig differenzierbaren Lösung u von $L(u) = f$ in irgendeinem in $\mathfrak{G}$ gelegenen Bereiche $\mathfrak{B}$ den bemerkenswerten Integralformeln[20])

$$\frac{\partial^2 u(X)}{\partial x_i \partial x_k} = J_1(X) + J_2(X) + J_{\mathfrak{R}}(X),$$

$$J_1(X) = -\int_{\mathfrak{B}}^{(n)} \{f(Y) - f(X)\} \frac{\partial^2 G}{\partial y_i \partial y_k}\, dY,$$

(3.1)

$$J_2(X) = \sum_{\nu,\mu} \int_{\mathfrak{B}}^{(n)} \{a_{\nu\mu}(Y) - a_{\nu\mu}(X)\} \frac{\partial^2 u}{\partial y_\nu \partial y_\mu} \frac{\partial^2 G}{\partial y_i \partial y_k}\, dY,$$

$$J_{\mathfrak{R}}(X) = \sum_{\nu\ \mu} a_{\nu\mu}(X) \int_{\mathfrak{R}}^{(n-1)} \left(\frac{\partial u}{\partial y_\mu} \frac{\partial^2 G}{\partial y_i \partial y_k} - u \frac{\partial^3 G}{\partial y_i \partial y_k \partial y_\mu} \right) d_\nu Y - f(X) \int_{\mathfrak{R}}^{(n-1)} \frac{\partial G}{\partial y_k}\, d_i Y$$

genügen, wo $G(X, Y)$ die durch (1.9) eingeführte Funktion ist. Im speziellen Falle $a_{\nu\nu} = 1$, $a_{\nu\mu} = 0$ für $\nu \neq \mu$ gehen diese Formeln direkt in die O. Hölderschen Formeln für die zweiten Ableitungen der Lösungen der Poissonschen Differentialgleichung über. Die einzige im allgemeinen Falle auftretende Komplizierung besteht, wie man sieht, in dem nochmaligen Auftreten der zweiten Ableitungen in den Integralen. Diesem Umstande sind, da wir im folgenden auf die sinngemäße Verallgemeinerung der Hölderschen Sätze hinauswollen, die Überlegungen von § 2 angepaßt.

Beweis von (3.1). Die Gleichung $L(u) = f$ werde in der Form

$$L^0(u) = \sum_{\nu,\mu} a_{\nu\mu}(X_0) \frac{\partial^2 u}{\partial x_\nu \partial x_\mu} = f(X) - \sum_{\nu,\mu} \{a_{\nu\mu}(X) - a_{\nu\mu}(X_0)\} \frac{\partial^2 u}{\partial x_\nu \partial x_\mu}$$

geschrieben. Wendet man auf $L^0(u)$, einen Differentialausdruck mit *konstanten* Koeffizienten, die Integralformel (1.10) an — dabei ist $G(X, Y)$ durch $\Gamma(a(X_0); Y - X)$ zu ersetzen —, so ergibt sich, da nach (1.8) $M^0(\Gamma) = 0$ ist,

$$u(X) = -\int_{\mathfrak{B}}^{(n)} \Gamma L_Y^0(u)\, dY + \sum_{\nu,\mu} a_{\nu\mu}(X_0) \int_{\mathfrak{R}}^{(n-1)} \left(\Gamma \frac{\partial u}{\partial y_\mu} - u \frac{\partial \Gamma}{\partial y_\mu} \right) d_\nu Y,$$

$$\Gamma = \Gamma(a(X_0); Y - X).$$

Bekanntlich dürfen die ersten Ableitungen $\frac{\partial u}{\partial X}$ in jedem inneren Punkte X

[20]) Sie setzen voraus, daß die Diskriminante $|a_{ik}|$ gleich Eins ist. Vgl. jedoch die Fußnote [15]).

von $\mathfrak{B}$ durch Differentiation unter dem Integralzeichen gebildet werden[21]), da die ersten Ableitungen $\frac{\partial \Gamma}{\partial X}$ für $Y = X$ nur von $(n-1)$-ter Ordnung unendlich werden. Nach Einsetzen des obigen Wertes von $L^0(u)$ ergibt sich also mit Rücksicht auf $\frac{\partial \Gamma}{\partial X} = -\frac{\partial \Gamma}{\partial Y}$

$$\frac{\partial u}{\partial x_i} = \int_{\mathfrak{B}}^{(n)} f(Y) \frac{\partial \Gamma}{\partial y_i} dY - \sum_{\nu,\mu} \int_{\mathfrak{B}}^{(n)} \{a_{\nu\mu}(Y) - a_{\nu\mu}(X_0)\} \frac{\partial^2 u}{\partial y_\nu \partial y_\mu} \frac{\partial \Gamma}{\partial y_i} dY$$
$$- \sum_{\nu,\mu} a_{\nu\mu}(X_0) \int_{\mathfrak{R}}^{(n-1)} \left(\frac{\partial \Gamma}{\partial y_i} \frac{\partial u}{\partial y_\mu} - u \frac{\partial^2 \Gamma}{\partial y_i \partial y_\mu} \right) d_\nu Y. \tag{3.2}$$

Wir bilden nunmehr $\frac{\partial^2 u}{\partial x_i \partial x_k}$. Auf das erste Integral der rechten Seite wende man den Hilfssatz 2 an. Wegen

$$\int_{\mathfrak{B}}^{(n)} \frac{\partial \Gamma}{\partial y_i} dY = \int_{\mathfrak{R}}^{(n-1)} \Gamma d_i Y$$

folgt $\left(\frac{\partial \Gamma}{\partial X} = -\frac{\partial \Gamma}{\partial Y} \right)$

$$\frac{\partial}{\partial x_k} \int_{\mathfrak{B}}^{(n)} f(Y) \frac{\partial \Gamma}{\partial y_i} dY = -\int_{\mathfrak{B}}^{(n)} \{f(Y) - f(X)\} \frac{\partial^2 \Gamma}{\partial y_i \partial y_k} dY - f(X) \int_{\mathfrak{R}}^{(n-1)} \frac{\partial \Gamma}{\partial y_k} d_i Y.$$

Die Differentiation der Randintegrale unter dem Integralzeichen ist natürlich gestattet. Bei den übrigen Integralen (J_2) ist dies nach dem Hilfssatz 1 wenigstens an der Stelle $X = X_0$ möglich (X_0 = innerer Punkt von $\mathfrak{B}$). Bildet man also in dieser Weise $\frac{\partial^2 u}{\partial x_i \partial x_k}$, so erhält man, wieder unter durchgängiger Benützung von $\frac{\partial \Gamma}{\partial X} = -\frac{\partial \Gamma}{\partial Y}$, endlich die Formeln (3.1), wenn man nachträglich wieder X statt X_0 schreibt. Dabei geht $\Gamma = \Gamma(a(X_0); Y - X)$ in die Funktion $G(X, Y)$ von (1.9) über.

Wir gehen nun zu ersten Anwendungen von (3.1) über und beweisen den

Satz 1. *Sind die Funktionen $a_{\nu\mu}(X)$ und $f(X)$ im Gebiete $\mathfrak{G}$ Hölder-stetig mit dem Exponenten α $(0 < \alpha < 1)$, so sind auch die zweiten Ableitungen jeder in $\mathfrak{G}$ zweimal stetig differenzierbaren Lösung von*

$$L(u) = f$$

in $\mathfrak{G}$ Hölder-stetig mit gleichem Exponenten α.

---

[21]) Vgl. den schönen Beweis von Erhard Schmidt, Schwarz-Festschrift, S. 365–383, insbes. § 1.

Beweis. Wir zeigen zunächst gemäß (3.1), daß $J_1(X)$, $J_2(X)$, $J_{\mathfrak{R}}(X)$ im Innern irgendeines in $\mathfrak{G}$ enthaltenen konvexen Bereiches $\mathfrak{B}$ überhaupt Hölder-stetig sind. Für $J_{\mathfrak{R}}(X)$ ist das sofort klar (der Exponent ist α), da die Integranden der Randintegrale Hölder-stetige Funktionen von X mit dem Exponenten α darstellen, und zwar gleichmäßig bezüglich Y, solange Y auf $\mathfrak{R}$ und X in einem ganz innerhalb $\mathfrak{B}$ gelegenen Teilbereiche variiert.

Auf $J_1(X)$, d. h. auf die Integrale

$$\int_{\mathfrak{B}}^{(n)} \{f(Y) - f(X)\} \frac{\partial^2 G}{\partial Y^2}\, dY \tag{3.3}$$

kann man den Hilfssatz 3 nebst Korollar mit

$$K(X, Z) = \frac{\partial \Gamma(a(X); Z)}{\partial Z}, \qquad \varrho(X) = f(X), \qquad \sigma(X) = 1$$

anwenden; denn mit Rücksicht auf (1.12) und die daran angeknüpften Bemerkungen sind alle Voraussetzungen über K erfüllt. Somit ist $J_1(X)$ im Innern von $\mathfrak{B}$ Hölder-stetig mit dem Exponenten α.

Somit bleiben noch $J_2(X)$, d. h. die Integrale

$$\int_{\mathfrak{B}}^{(n)} \{a(Y) - a(X)\} \frac{\partial^2 u}{\partial Y^2} \frac{\partial^2 G}{\partial Y^2}\, dY \tag{3.4}$$

zu untersuchen. Die Anwendung des Hilfssatzes 3 mit

$$K(X, Z) = \frac{\partial \Gamma(a(X); Z)}{\partial Z}, \qquad \varrho(X) = a(X), \qquad \sigma(X) = \frac{\partial^2 u}{\partial X^2}$$

ergibt für $J_2(X)$ wegen der bloßen Stetigkeit von σ nur die Hölder-Stetigkeit schlechthin. Da $\mathfrak{B}$ in $\mathfrak{G}$ beliebig war, folgt wegen (3.1) jedenfalls überhaupt die Hölder-Stetigkeit der zweiten Ableitungen $\frac{\partial^2 u}{\partial X^2}$ in $\mathfrak{G}$. Die Behauptung über den Exponenten ergibt sich nunmehr nachträglich, wenn man den Hilfssatz 3 nebst Korollar noch einmal heranzieht und die eben bewiesene Hölder-Stetigkeit von $\sigma(X)$ berücksichtigt.

Satz 2. *Sind die Funktionen $a_{\nu\mu}(X)$ und $f(X)$ im Gebiete $\mathfrak{G}$ mit Hölder-stetigen Ableitungen erster Ordnung versehen, so ist jede in $\mathfrak{G}$ zweimal stetig differenzierbare Lösung von*

$$L(u) = f$$

daselbst dreimal stetig differenzierbar.

Beweis. $\mathfrak{B}$ sei irgendein konvexer Bereich in $\mathfrak{G}$. Da die $a_{\nu\mu}(X)$ in $\mathfrak{G}$ einmal stetig differenzierbar sind, gilt natürlich dasselbe von der Randintegral-Summe $J_{\mathfrak{R}}(X)$ im Innern von $\mathfrak{B}$.

Bezüglich $J_1(X)$ schreibe man

$$\{f(Y)-f(X)\}\frac{\partial^2 G}{\partial y_i\,\partial y_k}=\frac{\partial}{\partial y_k}\Big[\{f(Y)-f(X)\}\frac{\partial G}{\partial y_i}\Big]-\frac{\partial f(Y)}{\partial y_k}\frac{\partial G}{\partial y_i}.$$

Somit ist

$$J_1(X)=\int\limits_{\mathfrak{B}}^{(n)}\frac{\partial f(Y)}{\partial y_k}\frac{\partial G}{\partial y_i}\,dY+\text{Randintegral},$$

wobei von dem Randintegral wieder das gleiche wie von $J_{\mathfrak{R}}(X)$ gilt. Auf das n-dimensionale Integral ist der Hilfssatz 2 mit

$$H(X,Y)=\frac{\partial G(X,Y)}{\partial y_i},\qquad \varrho(X)=\frac{\partial f(X)}{\partial x_k}$$

anwendbar, da alle Voraussetzungen wegen

$$\psi(X)=\int\limits_{\mathfrak{B}}^{(n)}H(X,Y)\,dY=\int\limits_{\mathfrak{R}}^{(n-1)}G(X,Y)\,d_iY$$

erfüllt sind. Also ist $J_1(X)$ innerhalb $\mathfrak{B}$ einmal stetig differenzierbar.

Bei $J_2(X)$, d. h. den Integralen (3.4) berücksichtige man die durch Satz 1 sichergestellte Hölder-Stetigkeit der $\frac{\partial^2 u}{\partial X^2}$. Auf diese Integrale ist abermals der Hilfssatz 2 mit

$$H(X,Y)=\{a(Y)-a(X)\}\frac{\partial^2 G}{\partial Y^2},\qquad \varrho(X)=\frac{\partial^2 u}{\partial X^2}$$

anwendbar. Die Voraussetzungen des Hilfssatzes sind erfüllt, da die Funktion

$$\psi(X)=\int\limits_{\mathfrak{B}}^{(n)}H(X,Y)\,dY=\int\limits_{\mathfrak{B}}^{(n)}\{a(Y)-a(X)\}\frac{\partial^2 G}{\partial Y^2}\,dY$$

aus dem gleichen Grunde wie $J_1(X)$ — vgl. den Ausdruck (3.3) — innerhalb $\mathfrak{B}$ einmal stetig differenzierbar ist. Somit besitzt auch $J_2(X)$ diese Eigenschaft, und $u(X)$ ist also nach (3.1) innerhalb $\mathfrak{B}$, und endlich in $\mathfrak{G}$ dreimal stetig differenzierbar.

Die wiederholte Anwendung der Sätze 1 und 2 liefert leicht den

Satz 3. *Sind im Gebiete $\mathfrak{G}$ die m-ten partiellen Ableitungen der Funktionen $a_{\nu\mu}(X)$, $b_\nu(X)$, $c(X)$, $f(X)$ vorhanden und Hölder-stetig mit dem Exponenten α, so ist jede in $\mathfrak{G}$ zweimal stetig differenzierbare Lösung von*

$$\sum_{\nu,\mu}a_{\nu\mu}(X)\frac{\partial^2 u}{\partial x_\nu\,\partial x_\mu}+\sum_\nu b_\nu(X)\frac{\partial u}{\partial x_\nu}+c(X)\,u=f(X)$$

daselbst $(m+2)$-mal Hölder-stetig — mit gleichem Exponenten α — differenzierbar.

Man beweist diesen Satz mit der gleichen Überlegung wie den Satz 4. *Die partielle Differentialgleichung zweiter Ordnung*

$$\Phi\left(\frac{\partial^2 u}{\partial x_1^2}, \frac{\partial^2 u}{\partial x_1 \partial x_2}, \ldots, \frac{\partial^2 u}{\partial x_n^2}; \frac{\partial u}{\partial x_1}, \ldots, \frac{\partial u}{\partial x_n}; u; x_1, \ldots, x_n\right) = 0$$

verhalte sich bezüglich einer vorgegebenen Lösung u an der Stelle $X = X_0$ elliptisch. Ist Φ als Funktion ihrer Argumente in der Umgebung der Stelle

$$\left(\frac{\partial^2 u}{\partial x_\nu \partial x_\mu}; \frac{\partial u}{\partial x_\nu}; u; x_\nu\right)_{X = X_0}$$

mit stetigen partiellen Ableitungen aller Ordnungen versehen, so gilt das gleiche von u als Funktion von X in der Umgebung von $X = X_0$, sobald u dort dreimal stetig differenzierbar ist.

Korollar. Hängt Φ von den zweiten Ableitungen $\frac{\partial^2 u}{\partial x_\nu \partial x_\mu}$ linear ab, so genügt die Voraussetzung, daß $u(X)$ zweimal stetig differenzierbar ist.

Beweis. Man schließe durch vollständige Induktion von der Existenz und Stetigkeit der $(m+2)$-ten Ableitungen auf die Existenz und Stetigkeit der $(m+3)$-ten Ableitungen von $u(X)$ $(m \geqq 1)$. Durch m-malige Differentiation der Gleichung $\Phi = 0$ nach den x_ν erhält man offenbar lineare elliptische Differentialgleichungen

$$\sum_{\nu,\mu} a_{\nu\mu} \frac{\partial^2 u_m}{\partial x_\nu \partial x_\mu} = \Phi_m$$

für die m-ten partiellen Ableitungen $u_m = \frac{\partial^m u}{\partial X^m}$, wobei die $a_{\nu\mu}$ in der Einleitung definiert wurden, und Φ_m nur die Ableitungen bis zur $(m+1)$-ten Ordnung von $u(X)$ enthält. Da die Existenz und Stetigkeit der u_{m+2} $(m \geqq 1!)$ schon als bewiesen vorausgesetzt war, sind die $a_{\nu\mu}$ und Φ_m als Funktionen von X einmal stetig differenzierbar, also nach Satz 1 die u_{m+2} Hölder-stetig. Daher sind nunmehr $a_{\nu\mu}$ und Φ_m als Funktionen von X einmal Hölder-stetig differenzierbar, also nach Satz 2 die u_m dreimal stetig differenzierbar, d. h. $u(X)$ ist $(m+3)$-mal stetig differenzierbar.

Diese Schlußweise tritt von $m = 1$ an in Kraft. Unter der Voraussetzung des Korollars kann man offenbar schon von $m = 0$ ab schließen.

§ 4.

Verschärfung des Satzes 4.

Für eine in einem Bereiche $\mathfrak{B}$ mit dem Exponenten α Hölder-stetige Funktion $\varphi(X)$ sei die Bezeichnung

$$H_\alpha(\varphi) = \text{Ob. Gr.} \, |\varphi(X'') - \varphi(X')| \cdot |X'' - X'|^{-\alpha} \qquad (X', X'' \text{ in } \mathfrak{B})$$

für die scharfe Hölder-Konstante eingeführt. Eine weitere Anwendung der Integralformeln (3.1) führt zu dem

Lemma 1. $a_{\nu\mu}(X)$ und $f(X)$ seien in einem Bereiche $\mathfrak{B}$ Hölderstetig mit dem Exponenten α $(0 < \alpha \leqq 1)$. Für die zweiten Ableitungen jeder in $\mathfrak{B}$ zweimal stetig differenzierbaren Lösung $u(X)$ der in $\mathfrak{B}$ elliptischen linearen Differentialgleichung

$$L(u) = \sum_{\nu,\mu} a_{\nu\mu}(X) \frac{\partial^2 u}{\partial x_\nu \partial x_\mu} = f(X)$$

gelten dann Ungleichungen der Form

$$\left|\frac{\partial^2 u}{\partial X^2}\right| < \text{konst.} \left[\frac{U}{d^2} + \frac{U'}{d} + F + d^\alpha H_\alpha(f)\right],$$

unter U, U', F die oberen Grenzen der u, $\frac{\partial u}{\partial X}$, f in $\mathfrak{B}$ verstanden, und wo d den Abstand des Punktes X vom Rande von $\mathfrak{B}$ bedeutet. Die Konstante kann ein für allemal fest gewählt werden, wenn die Funktionen $a_{\nu\mu}(X)$ hinreichend wenig abgeändert werden, jedoch so, daß dabei die Hölder-Konstanten $H_\alpha(a_{\nu\mu})$ unter einer festen Schranke bleiben[22]).

Beweis. Es bedeute K einen ganz im Innern von $\mathfrak{B}$ gelegenen Hyperkugel-Bereich, etwa mit dem Radius R. Wir operieren zunächst nur in K und verstehen unter δ den Abstand eines in K liegenden Punktes X vom Rande von K. Das Maximum

$$M = \text{Max}\, \delta^2 \left|\frac{\partial^2 u}{\partial X^2}\right| \qquad (X \text{ in } K), \tag{4.1}$$

genommen für alle zweiten Ableitungen von u, wird gewiß in einem inneren Punkte von K, den wir als Koordinatenursprung wählen wollen, angenommen. Ist δ_0 der Abstand dieses Punktes vom Rande von K, so ist

$$\delta_0 \leqq R \tag{4.2}$$

und nach (4.1)

$$\left|\frac{\partial^2 u}{\partial X^2}\right| \leqq \frac{M}{(|X| - \delta_0)^2}, \quad \left|\frac{\partial^2 u}{\partial X^2}\right|_{X=0} = \frac{M}{\delta_0^2} \qquad (|X| < \delta_0), \tag{4.3}$$

wobei die Gleichung so aufzufassen ist, daß das Gleichheitszeichen für mindestens eine der $\frac{\partial^2 u}{\partial x_\nu \partial x_\mu}$ zutrifft. Auf den neuen, gewiß ganz in K gelegenen Hyperkugel-Bereich $|X| < \vartheta \delta_0$ (mit noch zu bestimmendem $\vartheta < 1$) wende man nun die Formeln (3.1) an. Man beachte dabei im folgenden, daß bei hinreichend geringer Abänderung der $a_{\nu\mu}(X)$ die Kon-

[22]) Auf andere Anwendungen dieses Lemmas hoffe ich noch zurückzukommen. Von der oben angegebenen besonderen Form der Schranke wird hier kein Gebrauch gemacht.

stanten in den Ungleichungen (1.15) ein für allemal fest gewählt werden können. Dann folgt für den Punkt $X = 0$

$$(4.4) \qquad |J_1(0)| < \text{konst.}\, H_\alpha(f) \int\limits_{|Y|<\vartheta\delta_0}^{(n)} |Y|^{\alpha-n}\, dY = \text{konst.}\, H_\alpha(f)\,(\vartheta\delta_0)^\alpha.$$

Weiter erhält man, da wegen (4.3) $\left|\frac{\partial^2 u}{\partial Y^2}\right| < M\delta_0^{-2}(1-\vartheta)^{-2}$ für $|Y| < \vartheta\delta_0$ gilt,

$$(4.5) \qquad |J_2(0)| < \text{konst.}\, \frac{M}{\delta_0^2(1-\vartheta)^2} \cdot \int\limits_{|Y|<\vartheta\delta_0}^{(n)} |Y|^{\alpha-n}\, dY = \text{konst.}\, \frac{(\vartheta\delta_0)^\alpha M}{\delta_0^2(1-\vartheta)^2}.$$

Die Konstante hängt dabei nur von oberen Schranken für die $H_\alpha(a_{\nu\mu})$ ab. Für $J_\Re$, die Summe der über den Rand $|Y| = \vartheta\delta_0$ erstreckten Integrale, folgt durch leichte Anwendung von (1.4)

$$(4.6) \qquad |J_\Re(0)| < \text{konst.}\left(\frac{U}{(\vartheta\delta_0)^2} + \frac{U'}{\vartheta\delta_0} + F\right).$$

Wegen (4.3) und wegen (4.4), (4.5), (4.6) folgt dann

$$\left|\frac{\partial^2 u}{\partial X^2}\right|_{X=0} = \frac{M}{\delta_0^2} \leqq |J_1(0)| + |J_2(0)| + |J_\Re(0)|$$
$$< c \cdot \left[\frac{(\vartheta\delta_0)^\alpha}{\delta_0^2(1-\vartheta)^2} M + \frac{U}{(\vartheta\delta_0)^2} + \frac{U'}{\vartheta\delta_0} + F + (\vartheta\delta_0)^\alpha H_\alpha(f)\right],$$

wobei die Konstante c unter den Voraussetzungen des Lemmas fest gewählt werden kann. Nun ist $\delta_0 < D$, unter D den Durchmesser von $\mathfrak{B}$ verstanden, und daher

$$\frac{c(\vartheta\delta_0)^\alpha}{(1-\vartheta)^2} < cD^\alpha \cdot \frac{\vartheta^\alpha}{(1-\vartheta)^2}.$$

Wir erteilen dem $\vartheta < 1$ einen solchen festen Wert, daß hier der rechte Ausdruck $< \frac{1}{2}$ ausfällt. Dann ergibt sich wegen (4.2)

$$M < 2c\left[\frac{U}{\vartheta^2} + R\frac{U'}{\vartheta} + R^2 F + R^{2+\alpha}\vartheta^\alpha H_\alpha(f)\right].$$

Wegen (4.1) folgt also im Mittelpunkte von K ($\delta = R$)

$$\left|\frac{\partial^2 u}{\partial X^2}\right| \leqq \frac{M}{R^2} < \text{konst.}\left[\frac{U}{R^2} + \frac{U'}{R} + F + R^\alpha H_\alpha(f)\right].$$

Da nun dieser Mittelpunkt ein beliebiger innerer Punkt von $\mathfrak{B}$ sein darf, und R nur der Bedingung $R < d$ unterworfen ist, folgt endlich die gewünschte Ungleichung, wenn man $R \to d$ streben läßt.

Lemma 2. Die partielle Differentialgleichung

$$L^{(\delta)}(v) = \sum_{\nu,\mu} a_{\nu\mu}^{(\delta)}(X) \frac{\partial^2 v}{\partial x_\nu\, \partial x_\mu} = f^{(\delta)}(X)$$

sei für alle genügend kleinen $|\delta|$ im Bereiche $\mathfrak{B}$ elliptisch. Für diese $|\delta|$ seien die $a_{\nu\mu}^{(\delta)}$ und $f^{(\delta)}$ stetige Funktionen von X und δ, und die Hölder-Konstanten H_α dieser Funktionen mögen unterhalb einer von δ unabhängigen Schranke liegen. Ferner sei für jedes $\delta \neq 0$ eine Lösung v_δ vorgegeben. Konvergieren für $\delta \to 0$ v_δ und die $\frac{\partial v_\delta}{\partial X}$ gleichmäßig gegen stetige Funktionen v_0 und $\frac{\partial v_0}{\partial X}$, so ist v_0 innerhalb $\mathfrak{B}$ zweimal stetig differenzierbar.

Beweis. Nach dem Lemma 1 sind in jedem ganz innerhalb $\mathfrak{B}$ gelegenen Bereiche $\mathfrak{B}'$ die Funktionen $\left|\frac{\partial^2 v_\delta}{\partial X^2}\right|$ sicher unterhalb einer von δ unabhängigen Schranke gelegen. Man wende nun wiederum die Integralformeln (3.1) auf die Gleichung $L^{(\delta)}(v_\delta) = f^{(\delta)}$ im Bereiche $\mathfrak{B}'$ an; dabei ist $G = G_\delta = \Gamma(a_\delta(X); Y - X)$ zu setzen. Zunächst ist es klar, daß $J_{\mathfrak{R}}(X) = J_{\mathfrak{R}}^{(\delta)}(X)$ eine stetige Funktion von X und δ darstellt (mit Einschluß von $\delta = 0$). Was nun J_1 und J_2 betrifft, so bemerken wir, daß die Integranden die Form

$$K_\delta(X, Y)\, \varphi_\delta(Y)$$

besitzen, wobei $|\varphi_\delta(Y)|$ in $\mathfrak{B}'$ unterhalb einer von δ unabhängigen Schranke gelegen ist, und wo K_δ mit Rücksicht auf die Voraussetzungen des Lemmas eine für X, Y in $\mathfrak{B}$, $X \neq Y$, $|\delta| \leqq \delta_0$ bei passendem δ_0 stetige Funktion von X, Y und δ darstellt mit der Eigenschaft

$$|K_\delta(X, Y)| < \frac{\text{konst.}}{|Y - X|^{n-\alpha}};$$

die Konstante ist dabei von δ unabhängig. Auf Grund dieser Eigenschaften schließt man mit Hilfe einer wohlbekannten Stetigkeitsbetrachtung, daß das Integral

$$F_\delta(X) = \int_{\mathfrak{B}'}^{(n)} K_\delta(X, Y)\, \varphi_\delta(Y)\, dY$$

eine für alle δ mit $0 < |\delta| \leqq \delta_0$ gleichgradig stetige Funktion von X in $\mathfrak{B}'$ darstellt[23]). Da $J_1^{(\delta)}$ und $J_2^{(\delta)}$ von dieser Form sind, schließt man endlich, daß die Funktionen $\frac{\partial^2 v_\delta}{\partial X^2}$ bezüglich aller hinreichend kleinen $|\delta| > 0$ gleichgradig stetige Funktionen von X sind. Aus dieser Gleichgradigkeit folgt, wie eine leichte Überlegung lehrt, zunächst, daß die Limites

$$\lim_{\delta=0} \frac{\partial^2 v_\delta}{\partial X^2}$$

---

[23]) Es handelt sich um dieselbe (nur um eine geringfügige Gleichmäßigkeitsbetrachtung vermehrte) Überlegung, mit welcher man die Stetigkeit der ersten Ableitungen des Newtonschen Potentials mit beschränkter Belegungsdichte beweist. Vgl. etwa loc. cit. [21]).

existieren, und zwar mit gleichmäßiger Konvergenz im Innern von $\mathfrak{B}'$, sodann, daß diese Limites die notwendig stetigen zweiten Ableitungen von $v_0(X)$ darstellen.

Wir beweisen nun den

Satz 5. *Der Satz 4 gilt unter der geringeren Voraussetzung, daß die Lösung $u(X)$ in der Umgebung von $X = X_0$ zweimal Hölder-stetig differenzierbar ist.*

Beweis. Man setze

$$\frac{u(X+\delta)-u(X)}{\delta} = v_\delta(X), \tag{4.7}$$

unter δ den Zuwachs irgendeines x_ν verstanden. Die ersten bzw. zweiten Ableitungen von u seien kurz mit u' bzw. u'' bezeichnet. Ferner setze man

$$\begin{aligned} X_t &= X + t\,\delta, \\ u_t &= (1-t)\,u(X) \;+ t\,u(X+\delta), \\ u_t' &= (1-t)\,u'(X) \;+ t\,u'(X+\delta), \\ u_t'' &= (1-t)\,u''(X) + t\,u''(X+\delta). \end{aligned} \tag{4.8}$$

Schreibt man kurz $\Phi = \Phi(u'', u', u, X)$, so ist für $u = u(X)$ unter Anwendung einer bekannten elementaren Formel

$$\begin{aligned} 0 = \frac{\Phi(u_1'', u_1', u_1, X_1) - \Phi(u_0'', u_0', u_0, X_0)}{\delta} \qquad (4.9) \\ = \sum \frac{u_1''-u_0''}{\delta} \int_0^1 \frac{\partial \Phi}{\partial u''}(u_t'', u_t', u_t, X_t)\,dt \\ + \sum \frac{u_1'-u_0'}{\delta} \int_0^1 \frac{\partial \Phi}{\partial u'}(u_t'', u_t', u_t, X_t)\,dt \\ + \frac{u_1-u_0}{\delta} \int_0^1 \frac{\partial \Phi}{\partial u}(u_t'', u_t', u_t, X_t)\,dt \\ + \int_0^1 \frac{\partial \Phi}{\partial X}(u_t'', u_t', u_t, X_t)\,dt \end{aligned} \Bigg\} = -f^{(\delta)}(X).$$

Die erste Summe auf der rechten Seite lautet mit Rücksicht auf die in der Einleitung definierten $a_{\nu\mu}$ und wegen

$$\frac{u_1-u_0}{\delta} = v_\delta(X)$$

genauer

$$L^{(\delta)}(v_\delta) = \sum_{\nu,\mu} a_{\nu\mu}^{(\delta)}(X) \frac{\partial^2 v_\delta}{\partial x_\nu \partial x_\mu}, \qquad a_{\nu\mu}^{(\delta)}(X) = \int_0^1 a_{\nu\mu}(u_t'', u_t', u_t, X_t)\,dt.$$

Wegen der Hölder-Stetigkeit der $u''(X)$ in der Umgebung von $X = X_0$ sind die $a_{\nu\mu}^{(\delta)}(X)$, überhaupt die in (4.9) auftretenden Integrale offenbar ebenfalls Hölder-stetige Funktionen von X für alle hinreichend kleinen $|\delta|$; bei passendem $\alpha > 0$ werden sogar die auf eine passende Umgebung von $X = X_0$ bezogenen Hölder-Konstanten H_α dieser Funktionen unterhalb einer von δ unabhängigen Schranke liegen. Die Hölder-Stetigkeit der in (4.9) stehenden Faktoren

$$\frac{u_1' - u_0'}{\delta} = \frac{u'(X+\delta) - u'(X)}{\delta} = \int_0^1 u''(X + t\delta)\, dt,$$

$$\frac{u_1 - u_0}{\delta} = \frac{u(X+\delta) - u(X)}{\delta} = \int_0^1 u'(X + t\delta)\, dt$$

ist ebenfalls sofort ersichtlich, und von den Hölder-Konstanten gilt dasselbe. Somit genügen $a_{\nu\mu}^{(\delta)}(X)$ und das $f^{(\delta)}(X)$ von (4.9) allen Voraussetzungen von Lemma 2. Da wegen (4.7) für $\delta \to 0$

$$v_\delta(X) \to v_0(X) = u'(X), \qquad \frac{\partial v_\delta(X)}{\partial X} \to \frac{\partial v_0(X)}{\partial X} = u''(X)$$

gleichmäßig gilt, folgt also, daß die $u'(X)$ zweimal differenzierbar sind. Da somit $u(X)$ dreimal stetig differenzierbar ist, folgt die Behauptung aus dem Satz 4.

II. Kapitel.

Beweis des analytischen Charakters der Lösungen von $\Phi = 0$.

§ 5.

Ein System nichtlinearer Integralgleichungen für die Lösungen von $\Phi = 0$.

Die partielle Differentialgleichung

$$\Phi\left(\frac{\partial^2 u}{\partial x_1^2}, \frac{\partial^2 u}{\partial x_1 \partial x_2}, \ldots, \frac{\partial^2 u}{\partial x_n^2}; \frac{\partial u}{\partial x_1}, \ldots, \frac{\partial u}{\partial x_n}; u; x_1, \ldots, x_n\right) = 0$$

verhalte sich bezüglich der vorgegebenen Lösung $u(X)$ elliptisch und analytisch im Koordinatenursprung $X = 0$. Für den im folgenden zu erbringenden Beweis der Analytizität von $u(X)$ an der Stelle $X = 0$ setzen wir $u(X)$ als sechsmal stetig differenzierbar voraus. Nach den Resultaten des ersten Kapitels ist diese Voraussetzung von selbst erfüllt, sobald die $\frac{\partial^2 u}{\partial X^2}$ Hölder-stetig sind.

Wir schreiben zur Abkürzung $u' = \frac{\partial u}{\partial X}, \ldots, u'''' = \frac{\partial^4 u}{\partial X^4}$ für die partiellen Ableitungen von u bis zur vierten Ordnung und demgemäß $\Phi = \Phi(u'', u', u, X)$.

Es bedeutet keine Einschränkung der Allgemeinheit, wenn

$$(5.1) \qquad u(0) = u'(0) = u''(0) = u'''(0) = u''''(0) = 0$$

angenommen wird; sonst betrachte man statt u die Funktion

$$v = u - u(0) - \frac{u'(0)}{1!} X - \frac{u''(0)}{2!} X^2 - \frac{u'''(0)}{3!} X^3 - \frac{u''''(0)}{4!} X^4$$

(es ist wohl nicht mißzuverstehen, was damit angedeutet ist), welche, wie leicht zu sehen, ebenfalls einer elliptischen Differentialgleichung der obigen Art genügt. Differenziert man nun die Gleichung $\Phi = 0$ viermal nach den x_ν, so erhält man nach leichter Überlegung Differentialgleichungen der Form

$$(5.2) \qquad L(u'''') = \sum_{\nu,\mu} a_{\nu\mu} \frac{\partial^2 u''''}{\partial x_\nu \partial x_\mu} = \sum b \frac{\partial^5 u}{\partial X^5} + f \quad [24])$$

für jede der vierten Ableitungen u'''' mit (die $a_{\nu\mu}$ wurden bereits in der Einleitung definiert)

$$\begin{aligned} a_{\nu\mu} &= a_{\nu\mu}(u'', u', u, X), \\ b &= b(u''', u'', u', u, X), \\ f &= f(u'''', u''', u'', u', u, X), \end{aligned}$$

wobei diese Funktionen an der Stelle

$$u'''' = u''' = u'' = u' = u = X = 0$$

analytisch von ihren Argumenten abhängen. Daher läßt sich gewiß ein (im folgenden festgehaltener) Hyperkugel-Bereich

$$|X| \leqq R_0$$

so angeben, daß die Form $\sum a_{\nu\mu} z_\nu z_\mu$ für $u'' = u''(X), \ldots$ und $|X| \leqq R_0$ dauernd positiv definit ist und daß die $a_{\nu\mu}$, b, f als Funktionen von $u'''', \ldots, X$ für alle Wertesysteme $u'''' = u''''(X), \ldots$ und $|X| \leqq R_0$ analytisch sind.

Wir wenden auf (5.2) im Bereiche $|X| \leqq R_0$ die Integralformel (1.10) an, und zwar nicht mit der Funktion $G(X, Y)$ von (1.9), sondern mit der für die folgenden Zwecke geeigneteren Funktion

$$\begin{aligned} H(X, Y) &= G(Y, X) = \Gamma(a(Y); Y - X), \\ a(Y) &= a_{\nu\mu}(Y) = a_{\nu\mu}(u''(Y), u'(Y), u(Y), Y). \end{aligned}$$

---

[24]) D. h. die fünften Ableitungen kommen nur linear vor mit Koeffizienten, die von den u'''' nicht abhängen. Durch zweimalige Differentiation nach X erhält man zunächst

$$L(u'') = \Psi(u''', u'', u', u, X)$$

für jede der zweiten partiellen Ableitungen u''. Differenziert man noch einmal, so folgt

$$L(u''') = \sum u'''' \Psi_1(u''', u'', u', u, X) + \Psi_2(u''', u'', u', u, X),$$

woraus sich (5.2) ergibt.

Wir erhalten

$$(5.3)\quad u''''(X) = \int_{|Y|<R_0}^{(n)} \{u'''' M_Y(H) - Hf\}\, dY - \sum \int_{|Y|<R_0}^{(n)} Hb \frac{\partial^5 u}{\partial Y^5}\, dY + \sum_{\nu} \int_{|Y|=R_0}^{(n-1)} N_\nu \left(u'''', H, \frac{\partial u''''}{\partial Y}, \frac{\partial H}{\partial Y}\right) d_\nu Y.$$

Von den Randintegralen sei gleich bemerkt, daß sie für $|X| < R_0$ analytische Funktionen von X darstellen. Für $|Y| = R_0$ sind nämlich $H, \frac{\partial H}{\partial Y}$, also auch die Ausdrücke N_ν analytische Funktionen von X im Gebiete $|X| < R_0$, und zwar gleichmäßig in Y. Im zweiten Summanden rechts integriere man partiell nach dem Schema

$$Hb \frac{\partial^5 u}{\partial Y^5} = \frac{\partial (Hb u'''')}{\partial Y} - u'''' \frac{\partial (bH)}{\partial Y},$$

wo u'''' eine passende (mit dem u'''' links in (5.3) nicht notwendig übereinstimmende) vierte Ableitung bedeutet. Dann ergibt sich endlich

$$(5.4)\quad u''''(X) = \int_{|Y|<R_0}^{(n)} W\, dY + \varphi(X),$$

unter $\varphi(X)$ eine gewiß für $|X| < R_0$ analytische Funktion verstanden, mit

$$W = u''''(Y)\, M_Y(H) - Hf(u''''(Y), \ldots, Y) + \sum u''''(Y) \frac{d(bH)}{dY},$$

$$M_Y(H) = \sum_{\nu,\mu}' \frac{d^2(a_{\nu\mu} H)}{dy_\nu\, dy_\mu}.$$

Wir gebrauchen im Moment das Differentiationssymbol d, wo es sich um nur mittelbare Abhängigkeit von Y handelt; z. B. ist in verständlicher symbolischer Ausdrucksweise $(Z = Y - X)$

$$\frac{dH}{dY} = \frac{\partial \Gamma}{\partial a}\frac{da}{dY} + \frac{\partial \Gamma}{\partial Z}, \quad \frac{d^2H}{dY^2} = \frac{\partial^2 \Gamma}{\partial a^2}\left(\frac{da}{dY}\right)^2 + 2\frac{\partial^2 \Gamma}{\partial a\, \partial Z}\frac{da}{dY} + \frac{\partial^2 \Gamma}{\partial Z^2} + \frac{\partial \Gamma}{\partial a}\frac{d^2 a}{dY^2},$$

$$\frac{da}{dY} = \frac{\partial a}{\partial u''} u''' + \ldots + \frac{\partial a}{\partial Y}, \quad \frac{db}{dY} = \frac{\partial b}{\partial u''''} u'''' + \ldots + \frac{\partial b}{\partial Y},$$

$$\frac{d^2 a}{dY^2} = \frac{\partial a}{\partial u''} u'''' + \ldots + \frac{\partial^2 a}{\partial Y^2}.$$

Man erkennt hieraus, daß W sich als Linearkombination der Größen

$$(5.5)\quad \Gamma, \frac{\partial \Gamma}{\partial a}, \frac{\partial \Gamma}{\partial Z}, \frac{\partial^2 \Gamma}{\partial a^2}, \frac{\partial^2 \Gamma}{\partial a\, \partial Z} \quad (Z = X - Y)$$

— die $\frac{\partial^2 \Gamma}{\partial Z^2}$ fallen wegen (1.8) heraus — darstellt, deren Koeffizienten sich ganz und rational aus den u'''', u''', u'', u' und aus partiellen Ableitungen

der $a_{\nu\mu}$, b, f nach ihren unmittelbaren Argumenten zusammensetzen. Mit Rücksicht auf die Form (1.14) der Größen (5.5) erhält man W endlich in der Gestalt

$$(5.6)\qquad W(u'''', u''', u'', u', u, Y; Z) = \frac{Q_3(u'''', \ldots, u, Y; Z)}{\Omega(a; Z)^{\frac{n+2}{2}}},$$

$$a = a_{\nu\mu} = a_{\nu\mu}(u'', u', u, Y),$$

wo Q_3 ein Polynom der z_ν mit lauter Gliedern von mindestens drittem Grade darstellt, dessen Koeffizienten analytische Funktionen der $u'''', \ldots, u, Y$ für alle Wertesysteme

$$u'''' = u''''(Y), \ldots, u = u(Y), \quad Y \text{ mit } |Y| \leqq R_0$$

derselben sind.

Fügen wir die Relation $u(X) = \sum_\nu x_\nu \int_0^1 \frac{\partial u}{\partial x_\nu}(tX)\,dt$ (man berücksichtige (5.1)) und die analogen, die u' durch die u'' usw. ausdrückenden Relationen zu (5.4) hinzu, so erhalten wir endlich das in verständlicher symbolischer Form geschriebene Integralgleichungssystem

$$(5.7)\qquad u''''(X) = \int_{|Y|<R_0}^{(n)} W(u''''(X), \ldots, u(Y), Y;\ Y - X)\,dY + \varphi(X),$$

$$(5.8)\qquad \begin{cases} u'''(X) = \sum x \int_0^1 u''''(tX)\,dt, \quad u''(X) = \sum x \int_0^1 u'''(tX)\,dt, \\ u'(X) = \sum x \int_0^1 u''(tX)\,dt, \qquad u(X) = \sum x \int_0^1 u'(tX)\,dt, \end{cases}$$

wobei jedenfalls die Funktionen $\varphi(X)$ gegebene analytische Funktionen in $|X| < R_0$ sind. Unser Ziel ist der Analytizitätsbeweis des vorgegebenen Lösungssystems $u(X), \ldots, u''''(X)$ von (5.7), (5.8) an der Stelle $X = 0$.

Für das Folgende sind noch ein paar Bemerkungen über die ersten Ableitungen von W nach seinen Argumenten von Nutzen. Man überzeugt sich ähnlich wie am Schluß von § 1, daß sie die Form

$$(5.9)\qquad \frac{\partial W}{\partial u^{(i)}}, \frac{\partial W}{\partial Y} = \frac{Q_5(u'''', \ldots, u, Y; Z)}{\Omega(a; Z)^{\frac{n+4}{2}}}; \quad \frac{\partial W}{\partial Z} = \frac{Q_4(u'''', \ldots, u, Y; Z)}{\Omega(a; Z)^{\frac{n+4}{2}}}$$

besitzen, wo Q_4 bzw. Q_5 Polynome der z_ν mit Gliedern vom Mindestgrad vier bzw. fünf und mit Koeffizienten derselben Natur wie oben bei Q_3 darstellen. W, $\frac{\partial W}{\partial u^{(i)}}$, $\frac{\partial W}{\partial Y}$ werden also für $Z = 0$ von $(n-1)$-ter, $\frac{\partial W}{\partial Z}$ von n-ter Ordnung unendlich.

§ 6.

Vorbereitungen zum Übergang ins Komplexe.

Aus Durchsichtigkeitsgründen soll im folgenden statt des Integralgleichungssystems (5.7), (5.8) das einfachere System

$$\begin{aligned} u(X) &= \int_{|Y|<R_0}^{(n)} W(u(Y), v(Y), Y; Y-X)\, dY + \varphi(X), \\ v(X) &= x_1 \int_0^1 u(tX)\, dt \end{aligned} \tag{6.1}$$

für nur zwei Funktionen u, v betrachtet werden. $\varphi(X)$ sei wie vorhin eine für $|X| < R_0$ analytische Funktion von X, und W sei wieder von der Gestalt

$$\begin{aligned} W(u, v, Y; Z) &= \frac{Q_3(u, v, Y; Z)}{\Omega(a, Z)^{\frac{n+2}{2}}}, \\ a_{\nu\mu} &= a_{\nu\mu}(u, v, Y). \end{aligned} \tag{6.2}$$

Die partiellen Ableitungen erster Ordnung sind dann von der Form

$$\frac{\partial W}{\partial u}, \frac{\partial W}{\partial v}, \frac{\partial W}{\partial Y} = \frac{Q_5(u, v, Y; Z)}{\Omega(a; Z)^{\frac{n+4}{2}}}, \quad \frac{\partial W}{\partial Z} = \frac{Q_4(u, v, Y; Z)}{\Omega(a; Z)^{\frac{n+4}{2}}}. \tag{6.3}$$

Vorgegeben seien zwei im Gebiete $|X| < R_0$ beschränkte, einmal stetig differenzierbare und den Gleichungen (6.1) genügende Funktionen $u(X), v(X)$ mit

$$u(0) = v(0) = 0. \tag{6.4}$$

Die $a_{\nu\mu}$, sowie die Koeffizienten der Polynome Q_3, Q_4, Q_5 seien als Funktionen ihrer Argumente u, v, Y an allen Stellen

$$u(Y), v(Y), Y \quad \text{mit} \quad |Y| < R_0$$

und deren Häufungsstellen analytisch, und es existiere ein $\omega > 0$, so daß

$$\begin{aligned} \Omega(a, Z) &\geqq \omega |Z|^2, \\ a_{\nu\mu} &= a_{\nu\mu}(u(Y), v(Y), Y), \ |Y| < R_0 \end{aligned} \tag{6.5}$$

gilt.

Wir beweisen, daß unter diesen Voraussetzungen $u(X), v(X)$ an der Stelle $X = 0$ analytisch sind. Man wird leicht bemerken, daß die folgenden auf das System (6.1) bezüglichen Überlegungen in genau derselben Weise auf das System (5.7), (5.8) anwendbar sind. Wir schreiben die Gleichungen (6.1) noch in der für das Folgende geeigneteren Form

$$\begin{aligned} u(X) &= \int_{|Y|<R}^{(n)} W(u(Y), v(Y), Y; X-Y)\, dY + \Phi(X), \\ v(X) &= x_1 \int_0^1 u(tX)\, dt \end{aligned} \tag{6.6}$$

mit

$$\Phi(X) = \int_{R<|Y|<R_0}^{(n)} W(u(Y), v(Y), Y; Y-X)\,dY + \varphi(X) \tag{6.7}$$

und beschränken den Punkt X auf das kleinere Gebiet $|X| < R$ mit noch zu bestimmendem $R < R_0$.

Von jetzt ab seien auch komplexe Werte für die Koordinaten der Punkte X, Y, Z usw. zugelassen, indem

$$X = X' + iX'',\ Y = Y' + iY'',\ Z = Z' + iZ''$$

geschrieben werden soll, wenn $x_\nu = x'_\nu + i x''_\nu$, $y_\nu = y'_\nu + i y''_\nu$, $z_\nu = z'_\nu + i z''_\nu$ ist. Unser nächstes Ziel besteht darin, dem Gleichungssystem (6.6) auch für komplexe Punkte X einen passenden Sinn zu erteilen. Zu diesem Zweck seien komplexe Umgebungen des reellen Gebietes $|X| < R_0$ von der Form

$$(R)_\gamma: \quad |X'| < R,\ |X''| < \gamma(R - |X'|)$$

$(\gamma > 0)$ eingeführt. Offenbar ist $(R)_\gamma$ konvex; denn mit X_1 und X_2 gehört auch der Punkt $X = tX_1 + (1-t)X_2$ zu $(R)_\gamma$, da

$$\begin{aligned} |X'| + \gamma|X''| &\leqq t|X'_1| + (1-t)|X'_2| + \gamma t|X''_1| + \gamma(1-t)|X''_2| \\ &\leqq \gamma t R + \gamma(1-t)R = \gamma R \end{aligned}$$

ist. Für γ sei von vornherein der durch

$$\frac{\omega}{3\gamma^2} = \text{Ob. Gr.} \sum_{\nu,\mu} |A_{\nu\mu}|, \tag{6.8}$$

$$A_{\nu\mu} = A_{\nu\mu}(u(Y), v(Y), Y),\ Y \text{ reell},\ |Y| < R_0$$

gegebene feste Wert zugrunde gelegt. (Es sei noch einmal daran erinnert, daß die $A_{\nu\mu}$ die algebraischen Komplemente der $a_{\nu\mu}$ bezeichnen.)

$\Phi(X)$ für komplexe X. Wir zeigen zunächst, daß die durch (6.7) bestimmte Funktion $\Phi(X)$ bei passender Wahl von R im komplexen Gebiete $(R)_\gamma$ regulär und beschränkt ist. Für den Realteil der Form $\Omega(a; Z)$ gilt nun offenbar

$$\Re[\Omega(a; Z)] = \sum_{\nu,\mu} A_{\nu\mu}(z'_\nu z'_\mu - z''_\nu z''_\mu), \tag{6.9}$$

$$a_{\nu\mu} = a_{\nu\mu}(u(Y), v(Y), Y),\ Y \text{ reell},\ |Y| < R_0,$$

also wegen $|z''_\nu| \leqq |Z''|$ und wegen (6.5) und (6.8)

$$\Re[\Omega] \geqq \omega\left(|Z'|^2 - \frac{1}{3\gamma^2}|Z''|^2\right).$$

Für $Z = Y - X$ mit irgendeinem X aus $(R)_\gamma$ und reellem Y, $|Y| > R$, ist nun

$$|Z'| = |Y - X'|,\ |Z''| = |X''| < \gamma(R - |X'|) < \gamma|Y - X'|, \tag{6.10}$$

und daher unter der Bedingung (6.9)

$$\Re[\Omega] \geqq \frac{2}{3}\omega|Y - X'|^2. \tag{6.11}$$

$\Omega(a; Z)$ mit (6.9) ist daher für $|Y| > R$ eine in $(R)_\gamma$ reguläre Funktion von X mit positivem Realteil, $\sqrt{\Omega}$ also — es ist die für reelle X positive Wurzel gemeint — eine in $(R)_\gamma$ eindeutige reguläre Funktion von X. Unter den Bedingungen (6.9) und für $|Y| > R$ ist also auch $W(u(Y), v(Y), Y, Y-X)$, der Integrand des Integrales in (6.7), eine in $(R)_\gamma$ eindeutige reguläre Funktion von X mit den Eigenschaften

$$(6.12) \qquad |W| < \frac{\text{konst.}}{|Y-X'|^{n-1}}, \quad \left|\frac{\partial W}{\partial Z}\right| < \frac{\text{konst.}}{|Y-X'|^{n}}; \quad Z = Y-X,$$

die man leicht aus (6.2), (6.3) und (6.11) — in den Polynomen Q benütze man die aus (6.10) folgenden Ungleichungen

$$|z_\nu| \leqq |Z'| + |Z''| < (1+\gamma)|Y-X'|$$ [25])

— ersieht. Da die Funktion $\varphi(X)$ von (6.7) im reellen Gebiete $|X| < R_0$ analytisch ist, kann man $\mathfrak{R}$ so klein wählen, daß $\varphi(X)$ in $(R)_\gamma$ eindeutig und regulär ist. Bei passender Wahl von R wird dann nach dem Obigen die Funktion $\Phi(X)$ von (6.7) in $(R)_\gamma$ regulär und außerdem beschränkt sein, da nach (6.12)

$$|\Phi(X)| < \text{konst.} \int\limits_{R<|Y|<R_0}^{(n)} \frac{dY}{|Y-X'|^{n-1}} + \text{konst.}; \quad X \text{ in } (R)_\gamma$$

ist, und das hier auftretende Integral bekanntlich beschränkt ist. Da weiterhin für reelle Y mit $|Y| > R$ und für $|X'| < R$ gewiß $|Y-X'|^{-n} < (R-|X'|)\,|Y-X'|^{-n+1}$ ist, folgt aus (6.7) und (6.12)

$$\left|\frac{\partial \Phi(X)}{\partial X}\right| < \frac{\text{konst.}}{R-|X'|} \int\limits_{R<|Y|<R_0}^{(n)} \frac{dY}{|Y-X'|^{n-1}} + \text{konst.}$$

Somit gilt bei hinreichend kleinem R

$$(6.13) \qquad |\Phi(X)| < \text{konst.}, \quad \left|\frac{\partial \Phi(X)}{\partial X}\right| < \frac{\text{konst.}}{R-|X'|}; \quad X \text{ in } (R)_\gamma$$

für Φ und ihre ersten Ableitungen $\frac{\partial \Phi}{\partial x_\nu}$.

Die Funktion W für komplexe Argumente u, v, Y, Z. Die Zahl $\varkappa > 0$ sei so gewählt, daß — man berücksichtige (6.4) — in dem komplexen Gebiete

$$(6.14) \qquad |u| < \varkappa, \quad |v| < \varkappa, \quad |Y'| < \varkappa, \quad |Y''| < \varkappa$$

die $a_{\nu\mu}$ (also auch die $A_{\nu\mu}$) sowie die Koeffizienten der Polynome Q in (6.2) und (6.3) reguläre und beschränkte Funktionen von u, v, Y sind,

[25]) Man hat $|Q_3| < \text{konst.}\,|Y-X'|^3$, $|Q_4| < \text{konst.}\,|Y-X'|^4$, $|Q_5| < \text{konst.}\,|Y-X'|^5$.

und daß im Gebiet (6.14)

$$\sum_{\nu,\mu} |A_{\nu\mu} - A^0_{\nu\mu}| < \frac{\omega}{3(1+\gamma)^2}, \tag{6.15}$$

$$A^0_{\nu\mu} = A_{\nu\mu}(0,0,0)$$

ausfällt. Z sei der Bedingung — ihre Bedeutung wird in § 7 klar hervortreten —

$$|Z''| < \gamma |Z'| \tag{6.16}$$

unterworfen. Wir betrachten W in dem durch (6.14) und (6.16) bestimmten komplexen Gebiete. Zunächst ist

$$\Omega(a; Z) = \sum_{\nu,\mu} A^0_{\nu\mu} z_\nu z_\mu + \sum_{\nu,\mu} (A_{\nu\mu} - A^0_{\nu\mu}) z_\nu z_\mu$$

und daher

$$\Re[\Omega] \geqq \sum_{\nu,\mu} A^0_{\nu\mu}(z'_\nu z'_\mu - z''_\nu z''_\mu) - \sum_{\nu,\mu} |A_{\nu\mu} - A^0_{\nu\mu}|\,|z_\nu|\,|z_\mu|.$$

Wegen $|z''_\nu| \leqq |Z''| < \gamma |Z'|$ und

$$|z_\nu| \leqq |z'_\nu| + |z''_\nu| \leqq |Z'| + |Z''| < (1+\gamma)|Z'| \tag{6.17}$$

folgt mit Rücksicht auf (6.5) $(Y=0)$

$$\Re[\Omega] > \{\omega - \gamma^2 \sum_{\nu,\mu} |A^0_{\nu\mu}| - (1+\gamma)^2 \sum_{\nu,\mu} |A_{\nu\mu} - A^0_{\nu\mu}|\}\,|Z'|^2,$$

und hieraus schließlich wegen (6.8) und (6.15)

$$\Re[\Omega] > \frac{\omega}{3}|Z'|^2 \tag{6.18}$$

im Gebiete (6.14), (6.16). Wie oben folgt hieraus, daß die für reelle Z positiv bestimmte Wurzel $\sqrt{\Omega}$ in diesem Gebiete eine eindeutige reguläre und von Null verschiedene Funktion von u, v, Y, Z darstellt. In diesem Gebiete ist also auch W eine reguläre Funktion ihrer Argumente u, v, Y, Z mit den aus (6.2), (6.3) und (6.17), (6.18) folgenden Eigenschaften

$$|W|, \left|\frac{\partial W}{\partial u}\right|, \left|\frac{\partial W}{\partial v}\right|, \left|\frac{\partial W}{\partial Y}\right| < \frac{M}{|Z'|^{n-1}}, \left|\frac{\partial W}{\partial Z}\right| < \frac{M^*}{|Z'|^n}; \tag{6.19}$$

die Zahlen M und M^* hängen hier zwar noch von einer oberen Schranke für $|Z'|$ ab (die Polynome Q sind nicht homogen), können indessen von vornherein festgelegt werden, da im folgenden stets $|Z'| < 2R < 2R_0$ sein wird.

§ 7.

Definition und Eigenschaften des Integrales in (6.6) für komplexe Punkte X.

Dem in der Fundamentalgleichung (6.6) auftretenden Integral

$$F(X) = \int_{|Y|<R}^{(n)} W(u(Y), v(Y), Y;\ Y - X)\, dY$$

soll nunmehr mit Hilfe einer schönen Idee von E. E. Levi auch für komplexe X ein Sinn erteilt werden. Zu diesem Zweck denke man sich die für reelle Y gegebenen Funktionen $u(Y)$, $v(Y)$ irgendwie stetig in das komplexe Gebiet $(R)_\gamma$ hinein fortgesetzt. Wir definieren dann ein ganz bestimmtes „Funktional", welches diesen stetigen Fortsetzungen u, v eine in $(R)_\gamma$ erklärte, für reelle X sich auf das obige (gegebene) $F(X)$ reduzierende Funktion von X in $(R)_\gamma$ zuordnet. Zur Motivierung der in § 8 anzustellenden Überlegungen sei bemerkt, daß dieses Funktional eine in $(R)_\gamma$ reguläre Funktion von X liefert, wenn u, v in $(R)_\gamma$ regulär sind. Infolge dieser Eigenschaft müßten $u(X)$ und $v(X)$, falls sie wirklich in $(R)_\gamma$ regulär sind, das „komplexe" Integralgleichungssystem (6.6), wo das Integral durch jenes Funktional ersetzt ist, wirklich befriedigen. Wir kehren den Gedankengang um, indem wir in § 8 das komplexe Integralgleichungssystem auflösen und die Regularität des Lösungssystems $u(X)$, $v(X)$ in $(R)_\gamma$ beweisen.

Ein wesentlicher Punkt bei der Fortsetzung des obigen Integrales für komplexe X besteht in der gleichzeitigen Verlegung des Integrationsgebietes in den komplexen Y-Raum. Ist X irgendein Punkt aus $(R)_\gamma$, so wählen wir als Integrationsgebiet die Mannigfaltigkeit $[X, R]$ aller Punkte Y, welche auf den offenen geradlinigen Verbindungsstrecken des Punktes X mit den Punkten P der reellen Hyperkugel $|Y| = R$ liegen. Wegen der Konvexität von $(R)_\gamma$ liegen mit X auch alle Punkte der Mannigfaltigkeit $[X, R]$ in $(R)_\gamma$. Eine Parameterdarstellung von $[X, R]$ ist sofort durch die Formeln

$$(7.1) \qquad \begin{gathered} Y = X + t(P - X); \quad 0 < t < 1, \\ P = P(\varphi_1, \varphi_2, \ldots, \varphi_{n-1}) \end{gathered}$$

gegeben, wobei $\varphi_1, \varphi_2, \ldots, \varphi_{n-1}$ Parameter auf der Hyperkugel

$$|P| = \sqrt{\sum p_\nu^2} = R$$

bedeuten. $\varphi_1, \ldots, \varphi_{n-1}, t$ sind dann die Parameter der n-dimensionalen Mannigfaltigkeit $[X, R]$. Das über die letztere erstreckte Integral über eine Funktion $f(Y)$ wird durch die Formel

$$(7.2) \qquad \int\limits_{[X,R]}^{(n)} f(Y)\, dY = \int\limits_{t=0}^{1} \int\limits_{(\varphi)}^{(n-1)} f(Y) \frac{\partial(y_1, \ldots, y_{n-1}, y_n)}{\partial(\varphi_1, \ldots, \varphi_{n-1}, t)} \prod_\nu d\varphi_\nu\, dt$$

definiert; (φ) bedeutet eine gewisse Punktmenge im $(n-1)$-dimensionalen Parameterraum, deren Punkten die Punkte der Hyperkugel umkehrbar eindeutig entsprechen. Die Funktionaldeterminante ergibt sich leicht in der Form

$$(7.3) \qquad \frac{\partial(y_1, \ldots, y_{n-1}, y_n)}{\partial(\varphi_1, \ldots, \varphi_{n-1}, t)} = t^{n-1} \sum_1^n g_\nu(P)(p_\nu - x_\nu) = t^{n-1} E(P, P - X),$$

wobei die $g_\nu(P)$ die $(n-1)$-reihigen Determinanten der Funktionalmatrix $\left\| \frac{\partial p_k}{\partial \varphi_i} \right\|$ darstellen. Hieraus resultiert sogleich die geometrische Bedeutung des Linearausdrucks E; bis auf den Faktor

$$G(P) = \sqrt{\sum g_\nu^2(P)}$$

ist nämlich $E(P, P-X)$ bei reellem X gleich der Entfernung des Punktes X von derjenigen Hyperebene, welche die Hyperkugel $|P| = R$ im Punkte P berührt[26]). Hieraus folgen unmittelbar die Ungleichungen

$$(7.4)\quad G(P)(R - |X|) \leqq E(P, P-X) \leqq G(P)|P-X|;\ X \text{ reell},\ |X| \leqq R.$$

Allgemein ist

$$E(P, T) = \sum_\nu g_\nu(P) t_\nu \leqq G(P)|T|.$$

Für komplexe $X = X' + iX''$ aus $(R)_\gamma$ ist also

$$|E(P, P-X)| \leqq E(P, P-X') + |E(P, X'')|,$$

$$|E(P, X'')| \leqq G(P)|X''| < \gamma G(P)(R - |X'|) \leqq \gamma E(P, P-X'),$$

und daher

$$|E(P, P-X)| < (1+\gamma) E(P, P-X');\quad X \text{ in } (R)_\gamma.$$

Eine einfache Folge hieraus und aus (7.2), (7.3) ist die Ungleichung

$$(7.5)\qquad \left| \int_{[X,R]}^{(n)} f(Y)\, dY \right| \leqq (1+\gamma) \int_{|Y'|<R}^{(n)} |f(Y)|\, dY',$$

wobei im rechten Integral Y den durch seine reelle Projektion Y' eindeutig auf $[X, R]$ bestimmten Punkt bedeutet.

Wir wenden uns nun der genaueren Betrachtung des Integrales

$$(7.6)\qquad \begin{aligned} F(X) &= \int_{[X,R]}^{(n)} W(u(Y), v(Y), Y; Y-X)\, dY \\ &= \int_{t=0}^{1} \int_{(\varphi)}^{(n-1)} W(u(Y), v(Y), Y; Y-X)\, t^{n-1} E(P, P-X) \prod_\nu d\varphi_\nu\, dt, \\ &\qquad Y = X + t(P-X) \end{aligned}$$

zu und bemerken zunächst, daß sämtliche Argumentwerte u, v, Y, Z mit

$$(7.7)\qquad \begin{aligned} &|u| < \varkappa, \quad |v| < \varkappa, \quad |Y'| < \varkappa, \quad |Y''| < \varkappa; \\ &Z = Y - X, \quad X \text{ in } (R)_\gamma, \quad Y \text{ auf } [X, R] \end{aligned}$$

in das Regularitätsgebiet (6.14), (6.16) der Funktion $W(u, v, Y; Z)$ hineinfallen; denn es ist wegen (7.1)

$$|Z''| = t|X''| < t\gamma(R - |X'|) \leqq t\gamma|P - X'| = \gamma|Z'|.$$

---

[26]) Man kann die φ_ν immer so wählen, daß $E(P, P-X') \geqq 0$ für $|X'| \leqq R$ ausfällt.

Wegen (6.19) — man berücksichtige die Bemerkung am Schluß von § 6 — gilt ferner, sobald (7.7) erfüllt ist,

$$(7.8)\quad |W|,\ \left|\frac{\partial W}{\partial u}\right|,\ \left|\frac{\partial W}{\partial v}\right|,\ \left|\frac{\partial W}{\partial Y}\right| < \frac{M}{|Y'-X'|^{n-1}},\qquad \left|\frac{\partial W}{\partial Z}\right| < \frac{M^*}{|Y'-X'|^n}.$$

Damit das Integral (7.6) überhaupt einen Sinn hat, verlangen wir, daß die in $(R)_\gamma$ hinein stetig fortgesetzten $u(X)$, $v(X)$ dort ständig den Bedingungen $|u| < \varkappa$, $|v| < \varkappa$ genügen; ferner sei von vornherein

$$(7.9)\qquad R < \frac{\varkappa}{1+\gamma}$$

gewählt; denn dann ist sicher für alle Punkte Y aus $(R)_\gamma$ $|Y'| < R < \varkappa$, $|Y''| < \gamma R < \varkappa$.

Stetigkeit von $F(X)$ in $(R)_\gamma$. Denkt man sich $u(X)$, $v(X)$ unter Einhaltung der obigen Bedingungen in $(R)_\gamma$ hinein stetig fortgesetzt, so wird die durch (7.6) dargestellte Funktion $F(X)$ ebenfalls in $(R)_\gamma$ stetig sein. Denn der Integrand im rechten Integral ($Y = X + t(P - X)$) ist zunächst eine für

$$0 < t < 1, \qquad X \text{ in } (R)_\gamma$$

stetige Funktion von X, P, t, und es gilt wegen (7.8)

$$(7.10)\qquad t^{n-1}|W| < \frac{M}{|P-X'|^{n-1}} \leqq \frac{M}{(R-|X'|)^{n-1}}.$$

Bezeichnet man daher mit $F_\varepsilon(X)$ das aus (7.6) entstehende Integral, wenn man nur über $\varepsilon < t < 1-\varepsilon$ integriert, so ist $F_\varepsilon(X)$ gewiß in $(R)_\gamma$ stetig. Somit ist auch $F = F_0$ in $(R)_\gamma$ stetig, da wegen (7.10) auf jedem abgeschlossenen Teil von $(R)_\gamma$ gleichmäßig $F_\varepsilon \to F_0$ konvergiert.

Differenzierbarkeit von $F(X)$ in $(R)_\gamma$. Die ins Komplexe fortgesetzten Funktionen $u(X)$, $v(X)$, $X = X' + iX''$ seien nun als einmal stetig differenzierbar nach X', X'' in $(R)_\gamma$ vorausgesetzt. Eine solche Fortsetzung ist jedenfalls möglich, da die für reelle X gegebenen Funktionen $u(X)$, $v(X)$ als einmal stetig differenzierbar vorausgesetzt waren[27]) (z. B. setze man $u(X) = u(X')$, $v(X) = v(X')$). Genügen nun die Ableitungen Bedingungen der Form

$$(7.11)\qquad \left|\frac{\partial u}{\partial X'}\right|,\ \left|\frac{\partial u}{\partial X''}\right|,\ \left|\frac{\partial v}{\partial X'}\right|,\ \left|\frac{\partial v}{\partial X''}\right| < \frac{\text{konst.}}{R-|X'|},\qquad X \text{ in } (R)_\gamma,$$

so ist $F(X)$ in $(R)_\gamma$ einmal stetig nach X', X'' differenzierbar, und es gelten Ungleichungen derselben Form für das Integral F

$$(7.12)\qquad \left|\frac{\partial F}{\partial X'}\right|,\ \left|\frac{\partial F}{\partial X''}\right| < \frac{\text{konst.}}{R-|X'|},\qquad X \text{ in } (R)_\gamma.$$

---

[27]) Im Original-Gleichungssystem (57), (58) war die Existenz und Stetigkeit der sechsten Ableitungen von u vorausgesetzt worden.

Zum Beweise überlegt man sich leicht, daß im rechten Integral in (7.6) unter dem Integralzeichen differenziert werden darf. Rein formal ergibt sich zunächst mit Rücksicht auf den analytischen Charakter von W als Funktion von u, v, Y, Z, d. h. wegen

$$\frac{\partial W}{\partial y_\lambda'} = -i\frac{\partial W}{\partial y_\lambda''} = \frac{\partial W}{\partial y_\lambda}, \qquad \frac{\partial W}{\partial z_\lambda'} = -i\frac{\partial W}{\partial z_\lambda''} = \frac{\partial W}{\partial z_\lambda} \qquad (i = \sqrt{-1}),$$

$$\frac{\partial F(X)}{\partial x_\lambda'} = J_1 + J_2 + J_3,$$

$$J_1 = \int_{t=0}^{1}\int_{(\varphi)}^{(n-1)} \left\{\frac{\partial W}{\partial u}\frac{\partial u}{\partial y_\lambda'} + \frac{\partial W}{\partial v}\frac{\partial v}{\partial y_\lambda'} + \frac{\partial W}{\partial y_\lambda}\right\}(1-t)\,t^{n-1}\,E(P, P-X)\prod_\nu d\varphi_\nu\,dt,$$

(7.13)

$$J_2 = -\int_{t=0}^{1}\int_{(\varphi)}^{(n-1)} \frac{\partial W}{\partial z_\lambda}\,t^n\,E(P, P-X)\prod_\nu d\varphi_\nu\,dt,$$

$$J_3 = -\int_{t=0}^{1}\int_{(\varphi)}^{(n-1)} W t^{n-1}\,g_\lambda(P)\prod_\nu d\varphi_\nu\,dt,$$

wobei unter den Integralzeichen $Y = X + t(P - X)$, $Z = Y - X$ eingesetzt zu denken ist. Statt dessen kann man die Integrale auch in der Form

$$J_1 = \int_{[X,R]}^{(n)} \left\{\frac{\partial W}{\partial u}\frac{\partial u}{\partial y_\lambda'} + \frac{\partial W}{\partial v}\frac{\partial v}{\partial y_\lambda'} + \frac{\partial W}{\partial y_\lambda}\right\}(1-t)\,dY,$$

(7.14)

$$J_2 = -\int_{[R,X]}^{(n)} \frac{\partial W}{\partial z_\lambda}\,t\,dY, \qquad J_3 = -\int_{[X,R]}^{(n)} W\frac{g_2(P)}{E(P, P-X)}\,dY$$

schreiben. Analoge Formeln erhält man natürlich für $\frac{\partial}{\partial x_\lambda''}$; bei diesen sind jedoch die Größen $\frac{\partial W}{\partial y_\lambda}$, $\frac{\partial W}{\partial z_\lambda}$ und $g_\lambda(P)$ in den Integranden von J_1, J_2 und J_3 mit dem Faktor $i = \sqrt{-1}$ zu multiplizieren. Hieraus folgt die wichtige Formel

(7.15)

$$\frac{\partial F(X)}{\partial x_\lambda'} + i\frac{\partial F(X)}{\partial x_\lambda''} = \int_{[X,R]}^{(n)} \left\{\frac{\partial W}{\partial u}\left(\frac{\partial u}{\partial y_\lambda'} + i\frac{\partial u}{\partial y_\lambda''}\right) + \frac{\partial W}{\partial v}\left(\frac{\partial v}{\partial y_\lambda'} + i\frac{\partial v}{\partial y_\lambda''}\right)\right\}(1-t)\,dY,$$

in welcher der analytische Charakter von W wesentlich benützt wird.

Die oben vorgenommene Differentiation unter dem Integralzeichen kann leicht gerechtfertigt werden; denn die Integranden in (7.13) sind für $0 < t < 1$, X in $(R)_\gamma$ wieder stetige Funktionen von X, P, t, welche außerdem — das geht aus sogleich anzugebenden Ungleichungen hervor — für alle X eines abgeschlossenen Teiles von $(R)_\gamma$, sowie für alle t und P beschränkt sind. Es führt dann eine der obigen analoge ε-Überlegung zum Ziel.

Aus (7.8) und (7.11) folgt nun leicht

$$(7.16)\quad \left|\frac{\partial W}{\partial u}\frac{\partial u}{\partial y_\lambda'}+\frac{\partial W}{\partial v}\frac{\partial v}{\partial y_\lambda'}+\frac{\partial W}{\partial y_\lambda}\right| < \frac{\text{konst.}}{R-|Y'|}\cdot\frac{1}{|Y'-X'|^{n-1}} = \frac{\text{konst.}}{R-|Y'|}\frac{t^{-n+1}}{|P-X'|^{n-1}},$$

$Y=X+t(P-X)$, $Z=Y-X$. Da weiter $Y'-(1-t)X'=tP$, und somit $|Y'|-(1-t)|X'|\leqq t|P|=tR=R-(1-t)R$ ist, folgt

$$(7.17)\qquad 1-t\leqq\frac{R-|Y'|}{R-|X'|}.$$

Ferner ist nach (7.8)

$$(7.18)\quad |W|<\frac{M}{|Y'-X'|^{n-1}}=\frac{Mt^{-n+1}}{|P-X'|^{n-1}},\qquad \left|\frac{\partial W}{\partial z_\lambda}\right|<\frac{M^*}{|Y'-X'|^n}=\frac{M^*t^{-n}}{|P-X'|^n}$$

und

$$(7.19)\qquad t=\frac{|Y'-X'|}{|P-X'|}\leqq\frac{|Y'-X'|}{R-|X'|}.$$

Außerdem ist für alle X aus $(R)_\gamma$

$$|E(P,P-X)|\geqq\Re[E]=E(P,P-X')\geqq G(P)(R-|X'|)$$

(vgl. (7.4)) und daher wegen $|g_\lambda|\leqq G$

$$(7.20)\qquad \left|\frac{g_\lambda(P)}{E(P,P-X)}\right|\leqq\frac{1}{R-|X'|}.$$

Für die absoluten Beträge der Integranden in (7.14) ergibt sich damit, und zwar für J_1 aus (7.16) und (7.17), für J_2 aus (7.18) und (7.19), für J_3 aus (7.18) und (7.20) eine gemeinsame obere Schranke der Form

$$\frac{\text{konst.}}{R-|X'|}\,\frac{1}{|Y'-X'|^{n-1}}.$$

Die Anwendung der Ungleichung (7.5) auf (7.14) ergibt also

$$\left|\frac{\partial F}{\partial x_\lambda'}\right|\leqq|J_1|+|J_2|+|J_3|<\frac{\text{konst.}}{R-|X'|}\cdot\int\limits_{|Y'|<R}^{(n)}\frac{dY}{|Y'-X'|^{n-1}}$$

und Entsprechendes für $\frac{\partial F}{\partial x_\lambda''}$, woraus schließlich wegen der Beschränktheit des hier auftretenden Integrales die Behauptung folgt.

§ 8.

Das komplexe Integralgleichungssystem. Existenz und Regularität der Lösung.

Die für reelle X gegebenen, den Gleichungen (6.6) genügenden Funktionen $u(X)$, $v(X)$ sollen nun so in das komplexe Gebiet $(R)_\gamma$ fortgesetzt

werden, daß sie für alle X aus $(R)_\gamma$ die beiden Integralgleichungen

$$u(X) = \int_{[X,R]}^{(n)} W(u(Y),\ v(Y),\ Y;\ Y-X)\,dY + \Phi(X), \tag{8.1}$$
$$v(X) = x_1 \int_0^1 u(tX)\,dt$$

befriedigen. Diese Fortsetzung kann in einfachster Weise durch sukzessive Approximationen konstruiert werden.

Wir definieren durch die Gleichungen

$$u_{\nu+1}(X) = \int_{[X,R]}^{(n)} W_\nu\,dY + \Phi(X), \qquad v_{\nu+1}(X) = x_1 \int_0^1 u_\nu(tX)\,dt, \tag{8.2}$$
$$W_\nu = W(u_\nu(Y),\ v_\nu(Y),\ Y;\ Y-X),$$

$$u_0(X) = u(X'), \qquad v_0(X) = v(X') \tag{8.3}$$

zwei Folgen von Funktionen $u_\nu(X)$, $v_\nu(X)$ in $(R)_\gamma$ und setzen

$$\varphi_\nu(X) = u_{\nu+1}(X) - u_\nu(X), \qquad \psi_\nu(X) = v_{\nu+1}(X) - v_\nu(X), \tag{8.4}$$

$$\sigma_\nu = \text{Ob. Gr.}\,(|\varphi_\nu(X)|, |\psi_\nu(X)|), \quad X \text{ in } (R)_\gamma. \tag{8.5}$$

Da sich $u_0(X)$, $v_0(X)$ für reelle X auf die gegebenen Funktionen $u(X)$, $v(X)$ reduzieren, wird dies auch von allen $u_\nu(X)$, $v_\nu(X)$ gelten. Wir müssen zunächst durch passende Wahl von R dafür sorgen, daß überhaupt alle Approximationen u_ν, v_ν gebildet werden können. Natürlich muß R von vornherein so klein gewählt werden, daß $\Phi(X)$ in $(R)_\gamma$ regulär ist und die Eigenschaften (6.13) besitzt, und daß alle $u_\nu(Y)$, $v_\nu(Y)$, Y in $(R)_\gamma$, die Grundbedingungen (7.7) erfüllen (vgl. (7.9)). Zu diesem Zweck sei zunächst R so klein angesetzt, daß — man berücksichtige (6.4) und (8.3) —

$$|u_0(X)|, \quad |v_0(X)| < \frac{\varkappa}{3}, \quad X \text{ in } (R)_\gamma \tag{8.6}$$

ausfällt. Da weiter wegen (8.2), (8.3), (8.4)

$$\varphi_0(X) = \{u_1(X) - \Phi(X)\} - \{u(X') - \Phi(X)\} = \int_{[X,R]}^{(n)} W_0\,dY - \int_{|Y'|<R}^{(n)} W_0\,dY'$$

ist, wird für genügend klein gewähltes R, wie man leicht einsieht,

$$\sigma_0 < \frac{\varkappa}{3} \tag{8.7}$$

ausfallen. Schließlich sei R so klein gewählt, daß außerdem noch für $|X'| < R$

$$2(1+\gamma)\,M \int_{|Y'|<R}^{(n)} \frac{dY}{|Y'-X'|^{n-1}} < \frac{1}{2}; \qquad (1+\gamma)\,R < \frac{1}{2} \tag{8.8}$$

ausfällt. Wir legen ein allen diesen Bedingungen genügendes R zugrunde und zeigen, daß dann die Approximationen u_ν, v_ν ständig in $(R)_\gamma$ den Bedingungen $|u|, |v| < \varkappa$ genügen und gleichmäßig in $(R)_\gamma$ gegen stetige Grenzfunktionen u, v konvergieren. Aus (8.2) und (8.4) folgt zunächst

$$(8.9)\quad \varphi_{\nu+1}(X) = \int_{[X,R]}^{(n)} (W_{\nu+1} - W_\nu)\, dY, \qquad \psi_{\nu+1}(X) = x_1 \int_0^1 \varphi_\nu(tX)\, dt.$$

Nehmen wir an, wir hätten schon bewiesen, daß $u_0, v_0, \ldots, u_{\nu+1}, v_{\nu+1}$ in $(R)_\gamma$ den Bedingungen $|u|, |v| < \varkappa$ genügen [28]). Wir schließen dann, daß dies auch für $u_{\nu+2}, v_{\nu+2}$ gilt. Zunächst ist nach einer bekannten Formel

$$W_{\nu+1} - W_\nu = \varphi_\nu(Y) \int_0^1 \frac{\partial W}{\partial u}\, dt + \psi_\nu(Y) \int_0^1 \frac{\partial W}{\partial v}\, dt,$$

wobei in den Integralen für die Argumente u, v

$$u = t u_{\nu+1}(Y) + (1-t) u_\nu(Y), \qquad v = t v_{\nu+1}(Y) + (1-t) v_\nu(Y)$$

eingesetzt zu denken ist. Diese Argumentwerte genügen sicherlich den Bedingungen $|u|, |v| < \varkappa$, und man hat wegen (7.8) und (8.5)

$$|W_{\nu+1} - W_\nu| \leqq \frac{2M\sigma_\nu}{|Y' - X'|^{n-1}},$$

somit unter Benutzung der Ungleichung (7.5) und wegen (8.8)

$$|\varphi_{\nu+1}(X)| \leqq 2(1+\gamma) M\sigma_\nu \int_{|Y'|<R}^{(n)} \frac{dY}{|Y' - X'|^{n-1}} < \frac{1}{2}\sigma_\nu.$$

Ferner ist wegen (8.9) und $|x_1| \leqq |X'| + |X''| < R + \gamma R$ mit Rücksicht auf (8.8)

$$|\psi_{\nu+1}(X)| \leqq (1+\gamma) R\sigma_\nu \leqq \frac{1}{2}\sigma_\nu,$$

also endlich $\sigma_{\nu+1} \leqq \frac{1}{2}\sigma_\nu$ und damit $\sigma_\mu \leqq 2^{-\mu}\sigma_0$ für $\mu \leqq \nu+1$. Dann folgt

$$\begin{aligned} |u_{\nu+2}(X)| = |u_0(X) + \varphi_0(X) + \ldots + \varphi_{\nu+1}(X)| &\leqq |u_0(X)| + \sigma_0 + \ldots + \sigma_{\nu+1} \\ &\leqq |u_0(X)| + 2\sigma_0 \end{aligned}$$

und Analoges für $v_{\nu+2}$, also wegen (8.6) und (8.7)

$$|u_{\nu+2}(X)|, \ |v_{\nu+2}(X)| < \frac{\varkappa}{3} + 2\frac{\varkappa}{3} = \varkappa.$$

Daher erfüllen sämtliche Approximationen für alle X in $(R)_\gamma$ die Bedingungen $|u|, |v| < \varkappa$, und gleichzeitig konvergieren wegen $\sigma_\nu \leqq 2^{-\nu}\sigma_0$ die

[28]) u_0, v_0, u_1, v_1 erfüllen wegen (8.6), (8.7), (8.4) und (8.5) gewiß diese Bedingungen.

$u_\nu(X)$, $v_\nu(X)$ in $(R)_\gamma$ gleichmäßig gegen stetige Grenzfunktionen $u(X)$, $v(X)$, welche natürlich die Gleichungen (8.1) befriedigen und für reelle X mit den gegebenen Funktionen übereinstimmen müssen.

Regularität von $u(X)$, $v(X)$ in $(R)_\gamma$. Wir beweisen dieselbe, indem wir noch einmal auf die sukzessiven Approximationen zurückgreifen und zeigen, daß $u_\nu(X)$, $v_\nu(X)$ für $\nu \to \infty$ sozusagen „immer regulärer" werden. Zunächst ist leicht zu beweisen, daß für alle ν stetige Ableitungen

$$\frac{\partial u_\nu}{\partial X'}, \quad \frac{\partial u_\nu}{\partial X''}, \quad \frac{\partial v_\nu}{\partial X'}, \quad \frac{\partial v_\nu}{\partial X''}$$

in $(R)_\gamma$ existieren und dem Betrage nach unter Schranken der Form

$$\frac{\text{konst.}}{R - |X'|}$$

gelegen sein müssen. Wegen (8.3) sind zunächst $u_0(X)$, $v_0(X)$ gewiß einmal stetig differenzierbar in $(R)_\gamma$, da $u(X)$ und $v(X)$ für reelle X als einmal stetig differenzierbar vorausgesetzt waren. Man hat daher nach (8.2)

$$\frac{\partial u_1(X)}{\partial X'} = \frac{\partial}{\partial X'} \int_{[X,R]}^{(n)} W_1 \, dY + \frac{\partial \Phi}{\partial X}$$

und Analoges für $\frac{\partial u_1}{\partial X''}$, wobei das Integral nach dem Differenzierbarkeitssatz von § 7 in $(R)_\gamma$ gewiß einmal stetig nach X', X'' derivierbar ist. Wegen der dort angegebenen Betragsschranken für die Ableitungen und wegen (6.13) folgt daher die Richtigkeit der Behauptung für $\nu = 1$. In der gleichen Weise kann man sukzessive für $\nu = 2, 3, \ldots$ weiterschließen.

Wir führen nun die Operatoren

$$\nabla_1 = \frac{\partial}{\partial x_1'} + i \frac{\partial}{\partial x_1''}, \quad \nabla \quad \frac{\partial}{\partial x_2'} + i \frac{\partial}{\partial x_2''}, \ldots, \nabla_n = \frac{\partial}{\partial x_n'} + i \frac{\partial}{\partial x_n''}$$

und das abkürzende Symbol

$$\nabla = \frac{\partial}{\partial X'} + i \frac{\partial}{\partial X''}$$

ein. Diese Operation darf auf alle $u_\nu(X)$, $v_\nu(X)$ in $(R)_\gamma$ angewendet werden, und die obere Grenze

$$(8.10) \quad d_\nu = \text{Ob. Gr.} \{(R - |X'|)|\nabla u_\nu(X)|, \ (R - |X'|)|\nabla v_\nu(X)|\},$$

genommen für alle X aus $(R)_\gamma$ und alle $\nabla = \nabla_1, \ldots, \nabla_n$, ist nach dem Obigen für jedes ν endlich. Da $\Phi(X)$ in $(R)_\gamma$ regulär ist, hat man $\nabla \Phi = 0$, und die Anwendung der Formel (7.15) auf die Gleichungen (8.3) ergibt in verständlicher Weise

$$(8.11)\qquad \nabla u_{\nu+1}(X) = \int\limits_{[X,R]}^{(n)} \left\{ \frac{\partial W}{\partial u} \nabla u_\nu(Y) + \frac{\partial W}{\partial v} \nabla v_\nu(Y) \right\} (1-t)\, dY,$$

$$(8.12)\qquad \nabla v_{\nu+1}(X) = x_1 \int\limits_0^1 t \nabla u_\nu(tX)\, dt,$$

wobei $\frac{\partial W}{\partial u}$, $\frac{\partial W}{\partial v}$ an den Stellen $u = u_\nu(Y)$, $v = v_\nu(Y)$ genommen sind. Für das $1-t$ in (8.11) gilt nach (7.17)

$$0 < 1-t \leqq \frac{R - |Y'|}{R - |X'|}.$$

Mit Rücksicht auf (7.8) und durch Anwendung von (7.5) folgt also wegen (8.10)

$$|\nabla u_{\nu+1}(X)| \leqq \frac{2(1+\gamma) M d_\nu}{R - |X'|} \int\limits_{|Y'|<R}^{(n)} \frac{dY}{|Y' - X'|^{n-1}}.$$

Ferner ist, da in $(R)_\gamma$ $|x_1| < (1+\gamma)\, R$ gilt,

$$|\nabla u_{\nu+1}(X)| \leqq (1+\gamma) R\, d_\nu \cdot \int\limits_0^1 \frac{dt}{R - t|X'|} \leqq \frac{(1+\gamma) R d_\nu}{R - |X'|}.$$

Wegen (8.8) ergibt sich daher $d_{\nu+1} \leqq \frac{1}{2} d_\nu$, somit $d_\nu \leqq 2^{-\nu} d_0$ und nach (8.10) für alle ν

$$(8.13)\qquad |\nabla u_\nu(X)|,\ |\nabla v_\nu(X)| \leqq \frac{d_0}{2^\nu (R - |X'|)},\quad X \text{ in } (R)_\gamma$$

für $\nabla = \nabla_1, \ldots, \nabla_n$. Hieraus läßt sich nun die Regularität der Funktionen $u(X) = \lim u_\nu(X)$, $v(X) = \lim v_\nu(X)$ in $(R)_\gamma$, und damit die Analytizität der für reelle X gegebenen Funktionen $u(X)$, $v(X)$ leicht folgendermaßen erschließen. Halten wir im Moment $x_2, \ldots, x_n$ fest und betrachten wir die $u_\nu(X)$, $v_\nu(X)$ als Funktionen von x_1 in einem Bereiche der komplexen x_1-Ebene. Es bedeute Δ irgendein nebst seinem Inneren in diesem Bereiche gelegenes Dreieck. Für das längs Δ über u_ν erstreckte Integral gilt dann

$$\left| \int\limits_\Delta u_\nu\, dx_1 \right| \leqq U^2 \cdot \text{Ob. Gr.} |\nabla_1 u_\nu|,\ ^{29})$$

---

[29]) Ist $f(x) = f(x' + ix'')$ eine in einem einfach zusammenhängenden Bereiche einmal stetig nach x', x'' differenzierbare Funktion, so gilt allgemein für jedes in diesem Bereiche gelegene Dreieck Δ (U = Umfang von Δ) die Ungleichung

$$\left| \int\limits_\Delta f(x)\, dx \right| \leqq U^2 \{\text{Ob. Or.} |\nabla f(x)|;\ x \text{ in } \Delta\}.$$

Zum Beweise kann man dieselbe Überlegung heranziehen, die man zum heute üblichen Beweise des Cauchyschen Integralsatzes anstellt (vgl. etwa Bieberbach, Funktionentheorie I (1921), § 6). Man berücksichtige dabei die leicht einzusehende Tatsache, daß $|\nabla f(x)|$ den Maximalbetrag der Schwankung der im Punkte x nach allen möglichen Richtungen $\overrightarrow{dx}$ gebildeten Ableitungen $\frac{df}{dx}$ darstellt.

unter U den Umfang von Δ verstanden; die obere Grenze ist für alle x_1 aus Δ zu nehmen. Wegen (8.13) konvergiert dieselbe für $\nu \to \infty$ gegen Null, und aus der Gleichmäßigkeit der Konvergenz in $u_\nu \to u$, $v_\nu \to v$ folgt

$$\int_\Delta u\, dx_1 = \int_\Delta v\, dx_1 = 0.$$

Bekanntlich folgt daraus das Verschwinden der Kurvenintegrale über jeden geschlossenen, nebst seinem Inneren in jenem x_1-Bereiche enthaltenen Weg der x_1-Ebene, nach dem Moreraschen Satze also die Regularität von u, v als Funktionen von x_1. Ebenso folgt natürlich die Regularität in $x_2, \ldots, x_n$ und damit die Regularität von $u(X), v(X)$ im Gebiete $(R)_\gamma$, w. z. b. w.

(Eingegangen am 9. Mai 1930.)

Über den Perronschen Integralbegriff und seine Beziehung zu den R-, L- und D-Integralen.

Von

J. Ridder in Baarn (Niederlande).

Das Ziel dieser Arbeit ist zweierlei.

An erster Stelle haben wir die Absicht, die bei der Perronschen Integraldefinition auftretenden Klassen von Ober- und Unterfunktionen so zu wählen, daß sie eine möglichst große Anzahl von Funktionen umfassen, ohne daß sich dabei der Umfang dieses Integralbegriffes ändere. Dazu werden die Begriffe „oberhalb totalstetig" und „unterhalb totalstetig" und Verallgemeinerungen davon eingeführt, welche auf ähnliche Weise aus dem Begriffe der Totalstetigkeit hervorgehen wie die Halbstetigkeit nach oben bzw. nach unten aus dem Begriffe der Stetigkeit. Während die in § 2 eingeführten Klassen von Ober- und Unterfunktionen und die Perron-Bauerschen Funktionenklassen[1]) teilweise übereinandergreifen, umfassen die in § 5 eingeführten Klassen alle vorigen Klassen. Die in § 6 gebildete Integraldefinition, welche auf der Anwendung jener allgemeineren Funktionenklassen beruht, ist nun umfassender als die Integraldefinition von Perron-Bauer (und als die ebenso weit führende Definition des speziellen Denjoyschen Integrals[2])); denn die mit ihrer Hilfe integrierbaren Funktionen sind diejenigen, welche auch integrierbar sind nach der Integraldefinition von Denjoy-Khintchine[3]). Die Hinzufügung einer einschränkenden Bedingung für die Funktionen einer jeden der beiden Klassen [die Ober- und Unterfunktionen von § 2 und die P.-B. Funktionen genügen diesen Bedingungen] liefert eine Integraldefinition (§ 8), die völlig äquivalent ist mit derjenigen von Perron-Bauer und bei der jedoch die Funktionenklassen *umfassender* sind.

[1]) Siehe Perron, Sitzungsber. Heidelberg Ak. Wiss. 1914, Math.-naturw. Klasse, Abt. A, 14. Abhandl., S. 1—16 und Bauer, Monatsh. Math. Phys. 26 (1915), S. 155 u. 186; auch Kamke, Das Lebesguesche Integral, S. 128 u. 132.

[2]) Siehe Denjoy, Comptes Rendus 154 (1912), S. 859—862; auch Hobson, Theory of functions I, Chapter VIII.

[3]) Siehe Khintchine, Comptes Rendus 162 (1916), S. 287—291.

Weitere einschränkende Bedingungen führen zu Integralbegriffen, welche äquivalent sind mit dem Riemannschen (§ 3) oder mit dem Lebesgueschen (§ 4 und Fußnote [24]).

Unser zweites Ziel war eine neue Verallgemeinerung der Klassen der Ober- und Unterfunktionen (§ 10), wodurch *die Perronsche Integraldefinition in eine Form gebracht werden konnte, die ebenso weit führt wie die Definition des allgemeinen Denjoyschen Integrals*[4]). Diejenige Form dieser Definition, welche sich am meisten an die P.-B. Integraldefinition anschließt, wobei jedoch die benutzten Funktionenklassen etwas eingeschränkt sind, ist die folgende:

$f(x)$ sei eine in (a, b) definierte Funktion.

Eine für $a \leqq x \leqq b$ stetige Funktion $\psi(x)$ wird eine Oberfunktion sein, wenn sie den Bedingungen genügt: a) $\psi(a) = 0$; b) *es läßt sich (a, b) überdecken von höchstens abzählbar unendlich vielen, perfekten Mengen (E_j) und einer abzählbaren Menge M, derartig, daß auf jedem E_j die „in bezug auf dieses E_j" genommenen unteren Derivierten von $\psi(x) \neq -\infty$ und $\geqq f(x)$ sind.*

Ändert man „unteren" in „oberen", $-\infty$ in $+\infty$ und $\geqq$ in $\leqq$, so erhält man die Definition einer Unterfunktion.

$f(x)$ wird nun über (a, b) integrierbar sein, wenn die untere Schranke der Oberfunktionen im Punkte b und die obere Schranke aller Unterfunktionen in b einander gleich sind, und ihr gemeinsamer Wert soll dann das bestimmte Integral von $f(x)$ über (a, b) definieren[5]).

Wenn man in dieser Definition unter b) keine Zerlegung zuläßt, sondern die „allgemeinen" unteren (bzw. oberen) Derivierten benutzt, so erhält man die ursprüngliche P.-B. Definition zurück.

Erstes Kapitel.

§ 1.

Definition 1. Eine im Intervall $(a \leqq x \leqq b)$ endliche Funktion $f(x)$ wird *in (a, b) unterhalb totalstetig* sein, wenn stets die Summe der Funktionsdifferenzen

$$\sum_{(k)} \{f(b_k) - f(a_k)\}$$

in den Endpunkten a_k, b_k von endlich vielen, nicht übereinander greifenden Teilintervallen einen unteren Limes $\geqq 0$ hat, sobald die Längensumme der Teilintervalle gegen Null konvergiert: es wird dabei immer $a_k < b_k$ angenommen.

---

[4]) Siehe Denjoy, Ann. Éc. Norm. (3) **33** (1916), S. 127—222 und (3) **34** (1917), S. 181—238; auch Hobson, loc. cit. [2]).

[5]) Wir bemerken, daß die Definition keinerlei Meßbarkeitsbegriffe enthält.

Definition 2. Die im Intervall $(a \leqq x \leqq b)$ endliche Funktion $f(x)$ wird *in* (a, b) *oberhalb totalstetig* sein, wenn stets die Summe jener Funktionsdifferenzen einen oberen Limes $\leqq 0$ hat, sobald die Längensumme der Teilintervalle gegen Null konvergiert.

Eine beiden Definitionen genügende Funktion ist *totalstetig in* (a, b) im gewöhnlichen (Vitalischen) Sinne, und umgekehrt.

Satz I. *Eine stetige Funktion* $f(x)$, *deren untere rechte* [oder linke] *Derivierte in* $(a \leqq x \leqq b)$ *nach unten beschränkt ist, wird in* (a, b) *unterhalb totalstetig sein.*

Beweis. Da die Differenzenquotienten $\frac{f(d)-f(c)}{d-c}$ (wobei $a \leqq c < d \leqq b$) und die extremen Derivierten, nach einem Satze von Dini[6]), dieselbe untere Schranke haben im abgeschlossenen Intervall (a, b), ist der Satz ohnehin klar.

Satz II. *Eine stetige Funktion* $f(x)$, *deren obere rechte* [oder linke] *Derivierte in* $(a \leqq x \leqq b)$ *nach oben beschränkt ist, wird in* (a, b) *oberhalb totalstetig sein.*

Die Existenz von endlichen, nach Denjoy und Perron integrierbaren Ableitungen, deren primitive Funktionen nicht totalstetig sind, zeigt, daß in diesen Sätzen „beschränkt nach unten bzw. nach oben" nicht ersetzt werden darf durch „$\neq -\infty$ bzw. $\neq +\infty$" oder durch „endlich". Umgekehrt brauchen die unteren [oberen] Derivierten nicht immer von $-\infty$ $[+\infty]$ verschieden zu sein, wenn die stetige Funktion unterhalb [oberhalb] totalstetig ist.

Satz III. *Die unteren Derivierten einer in* (a, b) *endlichen und unterhalb totalstetigen Funktion* $f(x)$ *können nur auf einer Nullmenge* $-\infty$ *werden.*

Beweis. Sonst existierte eine Menge H mit äußerem Maße $m_a(H) > 0$, in deren Punkten eine der unteren Derivierten (z. B. die rechte untere, $D_+ f$) $= -\infty$ wäre. Es sei A eine willkürliche, positive Zahl und $H_1, H_2, \ldots, H_k, \ldots$ eine Folge von in H enthaltenen Mengen, deren äußere Maße mit zunehmendem k abnehmen und nach Null konvergieren, ohne jedoch identisch Null zu werden.

Zu jedem Punkte $x^{(k)}$ einer Menge H_k (k willkürlich, aber fest) gehört eine abzählbare Folge von Intervallen $(x^{(k)}, x_j^{(k)})$, welche sich mit zunehmendem j in $x^{(k)}$ zusammenziehen und für die

$$f(x_j^{(k)}) - f(x^{(k)}) < -\frac{A}{m_a(H_k)} \times (x_j^{(k)} - x^{(k)})$$

ist. Wenn $0 < p < 1$ ist, so läßt sich nach dem Überdeckungssatze von

---

[6]) Siehe Hobson, loc. cit. [2]), §§ 280, 281 oder Carathéodory, Vorles. über reelle Funktionen, § 476, Satz 5.

Vitali[7]) eine endliche Anzahl von nicht übereinandergreifenden Intervallen aus der Menge aller $(x^{(k)}, x_j^{(k)})$ wählen, deren Gesamtmaß zwischen $(1-p)\,m_a(H_k)$ und $(1+p)\,m_a(H_k)$ liegt und wobei dann die Summe der Funktionsdifferenzen in den Endpunkten dieser Intervalle

$$< -\frac{A}{m_a(H_k)} \times (1-p)\,m_a(H_k) \quad \text{oder} \quad < -(1-p)\cdot A$$

sein wird.

Mit zunehmendem k wird das Maß der immer endlichen Anzahl von Intervallen gegen Null konvergieren, ohne daß jedoch die Summe der Funktionsdifferenzen einen unteren Limes $\geqq 0$ hat. Die Funktion wäre dann nicht unterhalb totalstetig in (a, b).

Satz IV. *Die oberen Derivierten einer in (a, b) endlichen und oberhalb totalstetigen Funktion $f(x)$ können nur auf einer Nullmenge $+\infty$ werden.*

Der Beweis folgt durch Anwendung von Satz III auf die Funktion $-f(x)$.

Satz V. *Jede in (a, b) endliche und unterhalb oder oberhalb totalstetige Funktion hat fast überall in (a, b) eine endliche Ableitung und ist somit meßbar.*

Der Satz folgt aus Satz III oder IV und einem bekannten Satze von Denjoy-Saks über die extremen Derivierten endlicher Funktionen[8]).

Satz VI. *Eine in (a, b) endliche und unterhalb totalstetige Funktion $f(x)$ und eine in (a, b) endliche und oberhalb totalstetige Funktion $g(x)$ haben eine in (a, b) unterhalb totalstetige Differenz: $f(x) - g(x)$.*

§ 2.

Definition 3. $f(x)$ sei eine in (a, b) definierte Funktion. Eine für $a \leqq x \leqq b$ stetige Funktion $\psi_1(x)$ heißt eine *zu $f(x)$ im Intervall (a, b) adjungierte Oberfunktion*, wenn: a) $\psi_1(a) = 0$ ist; b) $\psi_1(x)$ in (a, b) unterhalb totalstetig ist; c) die beiden unteren Derivierten von $\psi_1(x)$ fast überall in $(a, b) \geqq f(x)$ sind.

Definition 4. Ebenso soll eine für $a \leqq x \leqq b$ stetige Funktion $\varphi_1(x)$ eine *zu $f(x)$ im Intervall (a, b) adjungierte Unterfunktion* sein, wenn: a) $\varphi_1(a) = 0$ ist; b) $\varphi_1(x)$ in (a, b) oberhalb totalstetig ist; c) die beiden oberen Derivierten von $\varphi_1(x)$ fast überall in $(a, b) \leqq f(x)$ sind.

---

[7]) Siehe Carathéodory, loc. cit. [6]), § 288, Satz 1.

[8]) Dieser lautet: Bei einer auf einer willkürlichen Menge E definierten, endlichwertigen Funktion sind, ausgenommen in den Punkten einer Nullmenge, in den übrigen Punkten nur die folgenden vier Fälle möglich: 1. $D^+ = D^- = +\infty$, $D_+ = D_- = -\infty$; 2. $D^+ = D_+ = D^- = D_- =$ endlich; 3. $D^+ = +\infty$, $D_- = -\infty$, $D_+ = D^- =$ endlich; 4. $D^- = +\infty$, $D_+ = -\infty$, $D_- = D^+ =$ endlich. Siehe Saks, Fund. Math. 5 (1924), S. 98—104.

Satz VII. *Die Differenz einer jeden Oberfunktion $\psi_1(x)$ und einer jeden Unterfunktion $\varphi_1(x)$, adjungiert zu der in (a, b) definierten Funktion $f(x)$, ist eine nicht abnehmende Funktion.*

Beweis. Bei willkürlich positivem ε läßt sich eine positive Zahl δ bestimmen, derartig, daß für jede Menge von endlich vielen, nicht übereinandergreifenden Intervallen (c_k, d_k), deren Gesamtmaß $< \delta$ ist, die zu $\omega_1(x) \equiv \psi_1(x) - \varphi_1(x)$ gehörige Summe

$$\sum_{(k)} \{\omega_1(d_k) - \omega_1(c_k)\} > -\varepsilon \tag{1}$$

sein wird (Satz VI).

Aus 3c) und 4c) folgt nach Satz V, daß $\omega_1(x)$ in einem maßgleichen Kerne K von (a, b) eine Ableitung $\geqq 0$ hat. Für jeden Punkt x von K gibt es eine Folge von Intervallen (x, x_j), welche sich mit zunehmendem j in x zusammenziehen und für die

$$\omega_1(x_j) - \omega_1(x) > -\varepsilon \cdot (x_j - x) \tag{2}$$

ist. Nun kann man, nach dem Überdeckungssatze von Vitali[7]), eine endliche Anzahl derartiger, nicht übereinander greifender Intervalle auswählen, deren Komplementärmenge ein Maß hat $< \delta$.

Aus (1) und (2) folgt demnach

$$\omega_1(b) - \omega_1(a) > -\varepsilon(b - a) - \varepsilon$$

oder, da ε willkürlich ist:

$$\omega_1(b) - \omega_1(a) \geqq 0.$$

Da dasselbe Beweisverfahren auch auf jedes Teilintervall von (a, b) angewandt werden kann, ist $\omega_1(x)$ nicht abnehmend in (a, b).

Definition 5. Zu der in (a, b) beschränkten Funktion $f(x)$ nimmt man den Wert des *unbestimmten oberen P-Integrals* in einem willkürlichen Punkte x von (a, b) gleich der unteren Schranke aller ψ_1-Werte in diesem Punkte, während das *unbestimmte untere P-Integral* in x definiert wird gleich der oberen Schranke aller φ_1-Werte in x. Wir schreiben

$$\Psi_1(x) \equiv \overline{\int_a^x} (P_1) f(x)\, dx = \operatorname{fin\,inf} \psi_1(x)$$

und

$$\Phi_1(x) \equiv \underline{\int_a^x} (P_1) f(x)\, dx = \operatorname{fin\,sup} \varphi_1(x).$$

Wenn in (a, b) $\Psi_1(x) = \Phi_1(x)$ ist, so definiert ihr gemeinsamer Wert *das unbestimmte P-Integral* $S_1(x)$ in (a, b).

Man sieht unmittelbar (Satz VII), daß immer $\Psi_1(x) \geqq \Phi_1(x)$ sein muß.

Satz VIII. *Wenn das bestimmte P-Integral über* (a, b) *existiert, d. h.* $\Psi_1(b) = \Phi_1(b)$ *ist, so existiert in* (a, b) *auch das unbestimmte P-Integral* $S_1(x)$.

Der Satz ist unschwer, mit Hilfe des Satzes VII, zu beweisen.

Auch beweist man leicht:

Satz IX. *Wenn* $S_1(x)$ *in* (a, b) *existiert, so sind, für eine willkürliche Oberfunktion* $\psi_1(x)$ *und eine willkürliche Unterfunktion* $\varphi_1(x)$, *die Differenzen:* $\psi_1(x) - S_1(x)$ *und* $S_1(x) - \varphi_1(x)$ *in* (a, b) *nicht-abnehmende Funktionen.*

Satz X. *Das unbestimmte P-Integral* $S_1(x)$ *einer beschränkten Funktion* $f(x)$ *besitzt fast überall eine Ableitung* $= f(x)$.

Beweis[9]). Betrachten wir eine Folge von Oberfunktionen $\{\psi_1^{(k)}(x)\}$ und eine Folge von Unterfunktionen $\{\varphi_1^{(k)}(x)\}$, welche mit zunehmendem k (gleichmäßig) nach $S_1(x)$ konvergieren. In einem maßgleichen Kerne K von (a, b) hat jedes $\psi_1^{(k)}$ und jedes $\varphi_1^{(k)}$ eine endliche Ableitung (Satz V).

Es sei ξ ein Punkt von K. Dann wird, nach Satz IX, für ein willkürlich gewähltes k und jedes $x > \xi$:

$\psi_1^{(k)}(x) - S_1(x) \geqq \psi_1^{(k)}(\xi) - S_1(\xi)$ und $S_1(x) - \varphi_1^{(k)}(x) \geqq S_1(\xi) - \varphi_1^{(k)}(\xi)$ oder

$$\psi_1^{(k)}(x) - \psi_1^{(k)}(\xi) \geqq S_1(x) - S_1(\xi) \geqq \varphi_1^{(k)}(x) - \varphi_1^{(k)}(\xi)$$

sein. Aus diesen Ungleichungen und den damit übereinstimmenden, für den Fall $x < \xi$, folgt, daß die vier extremen Derivierten von $S_1(x)$ in ξ zwischen $\psi_1'^{(k)}(\xi)$ und $\varphi_1'^{(k)}(\xi)$ liegen. Also wird, nach dem Denjoyschen Satze[8]), $S_1'(x)$ existieren in einem maßgleichen Kerne K_0 von K.

Bei positivem ε sei $K_j(\varepsilon)$ die Menge derjenigen Punkte von K_0, in denen die Ableitungen von $\psi_1^{(j)}$ und $\varphi_1^{(j)}$ eine Differenz $\geqq \varepsilon$ haben. Das Maß dieser Menge wird mit zunehmendem j nach Null konvergieren. Sonst existierte eine positive Zahl δ, derartig, daß für unendlich viele Werte (j') $m[K_{j'}(\varepsilon)] \geqq \delta$ wäre. Dies ist jedoch unmöglich. Anwendung des Vitalischen Überdeckungssatzes[7]) würde zeigen, daß für die nicht-abnehmenden Funktionen $\{\psi_1^{(j')} - \varphi_1^{(j')}\}$:

$$\psi_1'^{(j')}(b) - \varphi_1'^{(j')}(b) \geqq \delta \cdot \varepsilon$$

wäre, was unmöglich ist, denn:

$$\lim_{j'=\infty} \{\psi_1^{(j')}(b) - \varphi_1^{(j')}(b)\} = 0.$$

Daraus folgt (siehe auch Def. 3 und 4), daß $S_1'(x)$ fast überall in (a, b) um weniger als ε von $f(x)$ abweicht. Da ε willkürlich ist, wird $S_1'(x)$ fast überall in (a, b) gleich $f(x)$ sein.

---

[9]) Der Grundgedanke dieses Beweises rührt von Saks her; siehe Saks, Mathematica **1** (1929), S. 122—124.

Satz A. *Die oben gegebene Definition des P-Integrals $S_1(x)$ beschränkter Funktionen ist äquivalent mit der Definition des L-Integrals dieser Funktionen.*

Beweis. $f(x)$ sei eine in (a, b) beschränkte und P_1-integrierbare Funktion. Dann wird, nach Satz IX, für $x_1 < x_2$, und wenn $\psi_1(x)$ und $\varphi_1(x)$ zwei adjungierte Ober- und Unterfunktionen sind:

$$(3) \qquad \psi_1(x_2) - \psi_1(x_1) \geqq S_1(x_2) - S_1(x_1) \geqq \varphi_1(x_2) - \varphi_1(x_1)$$

sein. Da f beschränkt ist, sind ψ_1 und φ_1 immer so zu wählen, daß ihre Differenzenquotienten nach oben bzw. nach unten beschränkt sind. Dadurch folgt aus (3), daß $S_1(x)$ in (a, b) totalstetig ist.

Außerdem hat $S_1(x)$ fast überall eine Ableitung $= f(x)$ (Satz X). Also wird

$$S_1(x) = \int_a^x (L)\, S_1'(x)\, dx = \int_a^x (L)\, f(x)\, dx$$

sein.

Wenn umgekehrt $f(x)$ integrierbar ist nach Lebesgue, so ist $\int_a^x (L)\, f(x)\, dx$ sowohl Ober- wie Unterfunktion (s. Satz I, II; Def. 3, 4), woraus die P_1-Integrierbarkeit unmittelbar folgt und außerdem, daß

$$\int_a^x (P_1)\, f(x)\, dx = \int_a^x (L)\, f(x)\, dx$$

ist.

§ 3.

Man kann sich fragen, ob man die Klassen der Ober- und Unterfunktionen nicht derartig einschränken könne, daß das P-Integral $S_1(x)$ mit dem R-Integral zusammenfallen würde. Die Antwort liefern die Definitionen 6 bis 8 und der Satz B.

Definition 6. $f(x)$ sei eine im abgeschlossenen Intervall (a, b), beschränkte Funktion. Eine in dem Intervall stetige Funktion $\psi_1^*(x)$ heißt eine *zu $f(x)$ in (a, b), im abgeänderten Sinne, adjungierte Oberfunktion*, wenn: a) $\psi_1^*(a) = 0$ ist; b) eine der unteren Derivierten von ψ_1^* in den Punkten des abgeschlossenen Intervalls $\geqq f(x)$ ist; c) diese untere Derivierte fast überall im Intervall nach oben halbstetig ist.

Aus dem Dinischen Satze[6]) folgt, daß alle Derivierten von ψ_1^* nach unten beschränkt sind, so daß, nach dem Satze von Denjoy[8]), fast überall in (a, b) eine endliche Ableitung von ψ_1^* existiert, welche $\geqq f(x)$ ist. Nach Satz I ist ψ_1^* dabei auch unterhalb totalstetig. Dadurch ist die Klasse der $\{\psi_1^*\}$ enthalten in der Klasse der durch Definition 3 eingeführten Funktionen $\{\psi_1\}$.

Definition 7. $f(x)$ sei beschränkt für $a \leqq x \leqq b$. Eine für diese ε-Werte stetige Funktion $\varphi_1^*(x)$ heißt eine *zu $f(x)$ in (a, b), im abgeänderten Sinne, adjungierte Unterfunktion*, wenn: a) $\varphi_1^*(a) = 0$ ist; b) eine der oberen Derivierten von φ_1^* in den Punkten des abgeschlossenen Intervalls $\leqq f(x)$ ist; c) diese obere Derivierte fast überall in (a, b) nach unten halbstetig ist.

Die $\{\varphi_1^*\}$ sind enthalten in der Klasse der $\{\varphi_1\}$ von Definition 4.

Definition 8. Ändert man in Definition 5 ψ_1, φ_1, Ψ_1, Φ_1, S_1 in ψ_1^*, φ_1^*, Ψ_1^*, Φ_1^*, S_1^*, so erhält man die Definition des *unbestimmten P-Integrals $S_1^*(x)$ im abgeänderten Sinne.*

Die Sätze VII bis X behalten ihre Gültigkeit bei den in Definition 8 gegebenen Abänderungen.

Satz B. *Die Definition 8 des P-Integrals $S_1^*(x)$ beschränkter Funktionen ist äquivalent mit der Definition des R-Integrals dieser Funktionen*[10]).

Beweis. Die für $a \leqq x \leqq b$ beschränkte Funktion $f(x)$ habe ein unbestimmtes Integral $S_1^*(x)$. Dann sind, nach dem Beweisverfahren des Satzes X, die Funktionen ψ_1^* und φ_1^* bei willkürlich positivem ε derartig zu wählen, daß für jedes ξ einer Menge M, deren Maß um weniger als ε von $b - a$ abweicht:

$$D_+ \psi_1^*(\xi) - D^- \varphi_1^*(\xi) < \varepsilon \tag{4}$$

ist; wir nehmen an, daß für die beiden ausgewählten Funktionen die rechte untere bzw. die linke obere Derivierte die in den Definitionen 6 und 7 genannten Halbstetigkeitseigenschaften haben.

Für ein zu einem maßgleichen Kerne K von M gehörendes ξ existiert ein positives δ, so daß bei $\xi - \delta < x < \xi + \delta$

$$D_+ \psi_1^*(x) < D_+ \psi_1^*(\xi) + \varepsilon \quad \text{und} \quad D^- \varphi_1^*(x) > D^- \varphi_1^*(\xi) - \varepsilon \tag{5}$$

ist. Aus (4), (5) und den Bedingungen 6b, 7b folgt für dieselben Werte von x:

$$|f(x) - f(\xi)| < 3\varepsilon.$$

Da ε willkürlich ist, wird $f(x)$ in einem maßgleichen Kerne von (a, b) stetig sein, also ein (R)-Integral haben.

Aus der Existenz von $S_1^*(x)$ folgt auch die Existenz von $S_1(x)$ und ihre Gleichheit. Nach Satz A ist nun

$$S_1^*(x) = \int_a^x (L) f(x)\, dx, \quad \text{also auch} = \int_a^x (R) f(x)\, dx.$$

[10]) Es ist nicht möglich, ohne die Gültigkeit des Satzes zu beeinträchtigen, die in den Bedingungen b) der Definitionen 6 und 7 enthaltenen Ungleichungen nur „fast überall" zu fordern, sogar nicht, wenn man dann ψ_1^* unterhalb totalstetig und φ_1^* oberhalb totalstetig in (a, b) annimmt. Das zeigt das Beispiel einer Funktion $f(x)$, die in (a, b) für alle irrationalen x-Werte $= 0$, für alle rationalen x-Werte $= 1$ ist.

Umgekehrt, wenn $f(x)$ R-integrierbar ist, so läßt sich aus $\int_a^x (R) f(x)\,dx$ eine Oberfunktion konstruieren im Sinne der Definition 6, welche in jedem Punkte von (a, b) beliebig wenig von $\int_a^x (R) f(x)\,dx$ abweicht.

Es sei A die obere Schranke von $|f(x)|$ im abgeschlossenen Intervall (a, b). Dann läßt sich eine abzählbare Folge von positiven, abnehmenden Zahlen $\eta_0 = 2A, \eta_1, \ldots, \eta_k, \ldots$ angeben, so daß $\sum_{(k)} \eta_k$ endlich sein wird. Die Menge H_k der Punkte von (a, b), in denen die Oszillation $\omega(x)$ von $f(x)$ den Ungleichungen

$$\eta_k \geqq \omega(x) > \eta_{k+1} \tag{6}$$

genügt, hat ein Jordansches Maß $= 0$ und läßt sich somit in eine endliche Anzahl von nicht übereinandergreifenden, offenen Intervallen $J_j^{(k)}$ einschließen, deren Randpunkte keine Verdichtungspunkte von H_k sind und deren Gesamtlänge δ_k beliebig klein ist. Es läßt sich eine für $a \leqq x \leqq b$ nichtnegative und stetige Funktion $G_k(x)$ konstruieren, welche $= 0$ ist in den nicht zu den $\{J_j^{(k)}\}$ gehörenden Punkten, $= \eta_k$ in denjenigen Punkten, wo (6) gilt, und $\leqq \eta_k$ in den übrigen Punkten von (a, b). Dann wird für jedes x von (a, b)

$$\int_a^x (R)\, G_k(x)\,dx < \eta_k \cdot \delta_k$$

sein.

Wenn jedes δ_k unterhalb der positiven Zahl Δ liegt, so wird in (a, b)

$$\sigma(x) \equiv \sum_{(k)} \int_a^x (R)\, G_k(x)\,dx < \Delta \cdot \sum_{(k)} \eta_k$$

und

$$\sum_{(k)} G_k(x) < \sum_{(k)} \eta_k$$

sein. Da also $\sum_{(k)} G_k(x)$ gleichmäßig konvergiert in (a, b), ist diese Summe stetig und gliedweise nach Riemann integrierbar. Daraus folgt, daß in jedem Punkte x des abgeschlossenen Intervalls (a, b)

$$\sigma(x) = \int_a^x (R) \sum_{(k)} G_k(x)\,dx$$

ist und eine Ableitung hat, welche $\geqq \omega(x)$ ist.

Da außerdem Δ willkürlich klein zu wählen ist, sieht man leicht, daß $\sigma(x) + \int_a^x (R) f(x)\,dx$ eine Oberfunktion der verlangten Art ist.

Auf gleiche Weise findet man eine Unterfunktion im Sinne der Definition 7, welche beliebig wenig vom unbestimmten R-Integral abweicht. Die Definition 8 ist anwendbar und $S_1^*(x)$ hat den gleichen Wert wie $\int_a^x (R) f(x)\,dx$.

Zweites Kapitel.

§ 4.

Auch bei nicht-beschränkten Funktionen $f(x)$ läßt sich das unbestimmte Integral $S_1(x)$ gemäß Definition 5 einführen. Die Sätze VIII bis X bleiben dabei gelten. Daneben läßt sich zeigen:

Satz C. *Die Definition* 5 *des unbestimmten P-Integrals* $S_1(x)$ *ist bei nicht-beschränkten Funktionen ebenso weittragend wie die Lebesguesche Integraldefinition* [11]).

Beweis. Wenn $S_1(x)$ existiert, so muß sie totalstetig sein in (a, b). Sonst existierte eine Zahl Δ und eine Folge von Mengen (M_k), welche jede für sich von endlich vielen, nicht übereinandergreifenden Intervallen $(x_j^{(k)}, x_{j+1}^{(k)})$ gebildet würden, so daß gleichzeitig

$$\lim_{k=\infty} \sum_{(j)} [S_1(x_{j+1}^{(k)}) - S_1(x_j^{(k)})] = \Delta \quad \text{und} \quad \lim_{k=\infty} m(M_k) = 0 \tag{7}$$

wäre.

Nehmen wir Δ positiv an; $\varphi_1(x)$ sei eine adjungierte Unterfunktion mit

$$S_1(b) - \varphi_1(b) < \frac{1}{2}\Delta.$$

Nach Satz IX ist $S_1(x) - \varphi_1(x)$ eine nicht-abnehmende Funktion. Daraus folgt bei willkürlich, aber fest gewähltem k:

$$\sum_{(j)} [S_1(x_{j+1}^{(k)}) - S_1(x_j^{(k)})] - \sum_{(j)} [\varphi_1(x_{j+1}^{(k)}) - \varphi_1(x_j^{(k)})] \leqq S_1(b) - \varphi_1(b) < \frac{1}{2}\Delta. \tag{8}$$

Aus (7) und (8) folgt:

$$\limsup_{k=\infty} \sum_{(j)} [\varphi_1(x_{j+1}^{(k)}) - \varphi_1(x_j^{(k)})] > \frac{1}{2}\Delta.$$

Dies ist jedoch unmöglich, denn φ_1 ist oberhalb totalstetig.

Auch die Annahme eines negativen Δ würde zu einem Widerspruch führen. $S_1(x)$ ist somit totalstetig in (a, b); sie hat fast überall eine Ableitung $= f(x)$ (Satz X). Also:

$$S_1(x) = \int_a^x (L) f(x)\,dx.$$

[11]) Man vergleiche dieses Resultat mit einem Satze von Hake; siehe Hake, Math. Annalen **83** (1921), § 1.

Wenn, umgekehrt, $f(x)$ L-integrierbar ist, so ist $\int_a^x (L) f(x)\,dx$ sowohl Oberfunktion ψ_1 wie Unterfunktion φ_1, also auch P-Integral im Sinne der Definition 5.

Drittes Kapitel.

§ 5.

Nach den Untersuchungen von Hake, Looman und Alexandroff[12]) ist die ursprüngliche Perron-Bauersche Integraldefinition ebenso weittragend wie die Definition des speziellen Denjoyschen Integrals. Wir ändern, wieder mit Hilfe von Totalstetigkeitseigenschaften, die Definition der adjungierten Funktionen derartig ab, daß ihre Klassen ausgedehnter werden als bei Perron und Bauer und als in den beiden vorigen Abschnitten. Hierdurch erhalten wir eine Integraldefinition, welche das spezielle Denjoysche Integral umfaßt und welche äquivalent ist mit der Definition des (allgemeineren) Denjoy-Khintchineschen Integrals.

Definition 9. Die im Intervall $(a \leqq x \leqq b)$ definierte Funktion $f(x)$ ist *auf der* in (a, b) liegenden *Menge E unterhalb totalstetig*, wenn stets die Summe der Funktionsdifferenzen

$$\sum_{(k)} \{f(b_k) - f(a_k)\}$$

in den zu E gehörenden Endpunkten a_k, b_k von endlich vielen, nicht übereinandergreifenden Intervallen einen unteren Limes $\geqq 0$ hat, sobald die Längensumme dieser Intervalle gegen Null konvergiert.

Definition 10. Änderung von „unteren Limes $\geqq 0$“ in „oberen Limes $\leqq 0$“ gibt die Bedingung, unter der $f(x)$ *auf E oberhalb totalstetig* sein soll.

Definition 10*. Eine auf E unterhalb totalstetige und oberhalb totalstetige Funktion wird eine *auf E totalstetige Funktion* genannt[13]).

Definition 11. $f(x)$ ist *in (a, b) unterhalb totalstetig im verallgemeinerten Sinne* (abgekürzt: UTV), wenn (a, b) sich in höchstens abzählbar unendlich viele, meßbare Teilmengen (H_k) zerlegen läßt, auf denen $f(x)$ unterhalb totalstetig sei[14]).

---

[12]) Siehe Hake, l. c.[11]), S. 119–142; Looman, Math. Annalen **93** (1924), S. 153–156; Alexandroff, Math. Zeitschr. **20** (1924), S. 213–222.

[13]) Vgl. z. B. Lebesgue, Leçons sur l'intégration, 2. Éd., S. 232, und Saks, Fund. Math. **15** (1930), S. 247.

[14]) Wir bemerken, daß ein einzelner Punkt als Menge betrachtet werden kann, auf der $f(x)$ unterhalb (oberhalb) totalstetig ist; dasselbe gilt für eine Menge von endlich vielen Punkten. Wenn E eine unendliche Menge ist, auf der $f(x)$ unterhalb (oberhalb) totalstetig ist, so können diejenigen Punkte von E, in denen $f(x)$ unendlich wird, keinen Verdichtungspunkt haben, wie aus der Definition 9 (10) folgt (vgl.

Fortsetzung der Fußnote [14]) auf nächster Seite.

Definition 12. $f(x)$ ist *in* (a, b) *oberhalb totalstetig im verallgemeinerten Sinne* (abgekürzt: OTV), wenn (a, b) sich in höchstens abzählbar unendlich viele, meßbare Teilmengen (H_k) zerlegen läßt, auf denen $f(x)$ oberhalb totalstetig sei [14]).

Definition 12*. Eine Funktion, welche UTV und OTV ist in (a, b), wollen wir nennen: *totalstetig auf* (a, b) *im verallgemeinerten Sinne* (abgekürzt: TV). Eine derartige Funktion hat die charakteristische Eigenschaft, daß sich (a, b) zerlegen läßt in eine endliche oder abzählbar unendliche Anzahl von meßbaren Mengen. auf denen die Funktion totalstetig ist.

Satz XI. *Jede Funktion, welche UTV oder OTV ist in* (a, b), *hat fast überall in* (a, b) *eine endliche approximative Ableitung* [15]) *und ist meßbar.*

Beweis. Es sei z. B. $f(x)$ UTV in (a, b) und E eine willkürliche aus den Mengen, auf denen $f(x)$ unterhalb totalstetig (abgekürzt: UT) ist. Durch Hinzufügung der Verdichtungspunkte von E, welche nicht zu E gehören, entsteht aus E eine abgeschlossene Menge E^*. Diejenigen Punkte von E^*, in denen unterer und oberer Limes von $f(x)$ in bezug auf E nicht beide endlich sind, können keinen Verdichtungspunkt haben. Wenn ein derartiger Verdichtungspunkt ξ existierte, so müßte dieser Punkt eine Umgebung $\mathfrak{U}(\xi)$ haben, welche nur Punkte von E enthält, in denen $f(x)$ unendlich ist, ξ selbst vielleicht ausgenommen. Denn sonst würde es eine Folge von Intervallen (u_n) geben, in deren beiden, zu E gehörenden Endpunkten $f(x)$ gleich $+\infty$ oder gleich $-\infty$ wäre und deren Längen mit zunehmendem n nach Null konvergierten; die Definition 11 hätte also keinen Sinn. Der Punkt ξ könnte sein: 1. Verdichtungspunkt von Punkten aus der Menge E^*, in denen der obere Limes von $f(x)$ auf $E = +\infty$ ist; 2. Verdichtungspunkt von Punkten von E^*, in denen der untere Limes von $f(x)$ auf $E = -\infty$ ist; oder 3. Verdichtungspunkt von beiden Gruppen von Punkten. In allen drei Fällen würde man bei willkürlich positivem A eine Folge von Intervallen (v_k) konstruieren können, deren Endpunkte zu E gehörten und mit zunehmendem k nach ξ konvergierten, während die Funktionsdifferenz in den Endpunkten eines jeden dieser Intervalle $< -A$ wäre. Die Funktion wäre somit nicht UT auf E.

---

den Beweis des Satzes XI); ihre Anzahl ist somit endlich. [Wenn E das abgeschlossene Intervall (a, b) ist, so kann $f(x)$ nur in den Punkten a und b unendlich werden.] Man sieht nun leicht, daß Funktionen, welche in (a, b) UTV (OTV) sind, in den Punkten einer höchstens abzählbar unendlichen Menge unendlich werden können.

[15]) Eine Zahl A heißt approximative Ableitung einer Funktion $f(x)$ im Punkte ξ, wenn für die Punkte (x') einer Menge $M(\xi)$, welche in ξ die Dichte 1 hat:

$$\lim_{x'=\xi} \frac{f(x') - f(\xi)}{x' - \xi} = A$$

ist.

Die Menge H der Punkte von E^*, in denen unterer und oberer Limes von $f(x)$ in bezug auf E nicht beide endlich sind, wird also von nur endlich vielen Punkten gebildet [16]).

Für jeden nicht zu H gehörenden Punkt von E^* muß eine Umgebung $\mathfrak{U}$ existieren, derartig, daß $\mathfrak{U} \cdot E^*$ keine Punkte von H enthält. Die Menge $E^* - H$ läßt sich, nach einem Satze von Lindelöf [17]), von einer abzählbaren Folge von derartigen Umgebungen $(\mathfrak{U}_k)$ überdecken. Definieren wir nun in jedem $\mathfrak{U}_k$ eine endliche Funktion $g_k(x)$, welche 1. gleich $f(x)$ ist auf $\mathfrak{U}_k \cdot E$; 2. in jedem Punkte ξ von $\mathfrak{U}_k \cdot (E^* - E)$ gleich einer Zahl ist, welche willkürlich zwischen den in bezug auf E genommenen (endlichen) unteren und oberen Limes von $f(x)$ in ξ liegt; 3. sich linear ändert in den zu $\mathfrak{U}_k - E^*$ komplementären, abgeschlossenen Intervallen von $\mathfrak{U}_k$; dann wird $g_k(x)$ unterhalb totalstetig sein in $\mathfrak{U}_k$. Nach Satz V hat sie fast überall in $\mathfrak{U}_k$ eine endliche Ableitung. Und da die Menge $\mathfrak{U}_k \cdot E$ in fast jedem ihrer Punkte die Dichte 1 hat, so wird $f(x)$ fast überall auf $\mathfrak{U}_k \cdot E$ und somit auch fast überall auf E eine endliche approximative Ableitung haben.

Da E eine willkürliche der Mengen ist, auf denen $f(x)$ UT ist, so wird $f(x)$ in einem maßgleichen Kerne von (a, b) eine endliche approximative Ableitung haben. Sie wird also auch fast überall in (a, b) approximativ stetig [18]) sein; daraus folgt, nach einem Satze von Stepanoff [19]), ihre Meßbarkeit in (a, b).

Satz XII. *Wenn zwei endliche Funktionen $f(x)$ und $g(x)$ UTV bzw. OTV sind in (a, b), so wird die Differenz $f(x) - g(x)$ UTV sein in (a, b).*

§ 6.

Hilfssatz. *Wenn sich das abgeschlossene Intervall (a, b) zerlegen läßt in endlich oder abzählbar unendlich viele Mengen (H_k), so hat jede in (a, b) enthaltene, perfekte Menge E ein Stück [20]), auf dem eine der Mengen (H_k) überall dicht liegt.*

---

[16]) Es sind für diese Punkte nur drei Fälle möglich: 1. ein derartiger Punkt ist linker Grenzpunkt von E und es existiert ein einziger Limes von $f(x)$ in bezug auf E in diesem Punkte, gleich $+\infty$; oder 2. der Punkt ist rechter Grenzpunkt von E und es existiert ein einziger Limes von $f(x)$ in bezug auf E in dem Punkte, gleich $-\infty$; oder 3. der Punkt ist isolierter Punkt von E mit $f = +\infty$ oder $-\infty$.

[17]) Siehe Carathéodory, loc. cit. [6]), § 60.

[18]) $f(x)$ heißt in ξ approximativ stetig, wenn für die Punkte (x') einer Menge $M(\xi)$, welche in ξ die Dichte 1 hat:

$$\lim_{x'=\xi} f(x') = f(\xi)$$

ist.

[19]) Siehe Stepanoff, Recueil Soc. Math. de Moscou **31** (1922—24), S. 487—489.

[20]) Unter einem Stück einer abgeschlossenen Menge E versteht man eine abgeschlossene, in einem abgeschlossenen Intervall (c, d) enthaltene Menge, die im Innern von (c, d) mit E zusammenfällt.

Beweis. Wäre dies nicht der Fall, so würde sich ein Stück $\tilde{\omega}_1$ von E bestimmen lassen, das keine Punkte von H_1 enthält; es existierte ein Stück $\tilde{\omega}_2$ von $\tilde{\omega}_1$, das keine Punkte von H_2 enthält; usw. Wenn die Anzahl der Mengen (H_k) endlich wäre (z. B. n), so existierte ein Stück $\tilde{\omega}_n$, das mit keiner der Mengen (H_k) Punkte gemeinsam hätte, was unmöglich ist. Ist die Anzahl der Mengen unendlich, so konvergieren die Endpunkte der $(\tilde{\omega}_k)$ zu den Randpunkten eines Grenzintervalls oder zu einem Grenzpunkt. Die zu E gehörenden Randpunkte dieses Intervalls oder der zu E gehörende Grenzpunkt könnten dann wieder zu keiner der Mengen (H_k) gehören und die Summe der Mengen (H_k) bildete also das gesamte Intervall nicht.

Definition 13. $f(x)$ sei eine in (a, b) definierte Funktion. Eine *zu* $f(x)$ *im Intervall* (a, b) *adjungierte Oberfunktion* $\psi_2(x)$ soll stetig sein für $a \leqq x \leqq b$ und daneben den folgenden Bedingungen genügen: a) $\psi_2(a) = 0$; b) $\psi_2(x)$ ist UTV in (a, b); c) die beiden unteren Derivierten von $\psi_2(x)$ sind fast überall in $(a, b) \geqq f(x)$.

Definition 14. Eine *zu* $f(x)$ *in* (a, b) *adjungierte Unterfunktion* $\varphi_2(x)$ ist stetig für $a \leqq x \leqq b$ und genügt den weiteren Bedingungen: a) $\varphi_2(a) = 0$; b) $\varphi_2(x)$ ist OTV in (a, b); c) die beiden oberen Derivierten von $\varphi_2(x)$ sind fast überall in $(a, b) \leqq f(x)$.

Satz XIII. *Die in den Definitionen* 13 *und* 14 *eingeführten Ober- und Unterfunktionen besitzen fast überall in* (a, b) *eine endliche Ableitung, wenn ihre Klassen gleichzeitig existieren.*

Das folgt aus den Bedingungen 13 c) und 14 c), unter Anwendung des Denjoyschen Satzes (loc. cit. [8])) [21]).

Satz XIV. *Die Differenz einer jeden Oberfunktion* $\psi_2(x)$ *und einer jeden Unterfunktion* $\varphi_2(x)$, *adjungiert zu der in* (a, b) *definierten Funktion* $f(x)$, *ist eine nicht-abnehmende Funktion.*

Beweis. Dieser verläuft mittels transfiniter Induktion.

Nach dem Hilfssatz enthält (a, b) eine nirgends dichte, abgeschlossene Menge E_1 mit der Eigenschaft, daß jeder Punkt ξ der in bezug auf (a, b) komplementären Menge $C(E_1)$ eine Umgebung $\mathfrak{U}(\xi)$ hat, in der eine der zu $\psi_2(x)$ gehörenden Mengen (H_k) (s. Def. 11) und eine der übereinstimmenden, zu $\varphi_2(x)$ gehörenden Mengen (s. Def. 12) überall dicht liegen; da $\psi_2(x)$ und $\varphi_2(x)$ außerdem stetig sind, so werden sie auch unterhalb bzw. oberhalb totalstetig sein in jener Umgebung.

---

[21]) Es zeigt sich daneben, daß eine Funktion, welche Ober- und Unterfunktionen hat, fast überall endlich sein muß.

Nach Satz VII wird $\omega_2(x) \equiv \psi_2(x) - \varphi_2(x)$ nicht-abnehmend sein in $\mathfrak{U}(\xi)$. Aus der Stetigkeit von $\omega_2(x)$ folgt jene Eigenschaft auch für jedes zu E_1 komplementäre, abgeschlossene Intervall von (a, b).

E_1 besteht aus dem perfekten Kerne K_1 und einem abzählbaren (in jedem Lückenintervall von K_1 reduziblen) Rest R_1[22]). Für ein zu K_1 komplementäres Intervall (a_k, b_k), das Punkte von R_1 enthält, ist $\omega_2(x)$ nicht-abnehmend in einem jeden abgeschlossenen Teilintervall U, das keine inneren Punkte von R_1 enthält, dann auch in jedem Lückenintervall der ersten Ableitung von R_1 und, sukzessive so weitergehend, in jedem Lückenintervall irgendeiner höheren Ableitung von R_1, also schließlich nach abzählbar vielen Operationen im abgeschlossenen Intervall (a_k, b_k). $\omega_2(x)$ ist nicht-abnehmend für jedes zu K_1 komplementäre, abgeschlossene Intervall von (a, b).

K_1 enthält (Hilfssatz) eine auf K_1 nirgends dichte abgeschlossene Menge E_2 mit der Eigenschaft, daß zu jedem Punkte ξ von $K_1 - E_2$ ein Stück $S(\xi)$ von K_1 existiert, dessen Endpunkte α und β ein ξ im Innern enthaltendes Intervall (α, β) bilden und auf dem $\psi_2(x)$ unterhalb und $\varphi_2(x)$ oberhalb totalstetig sind. Eine stetige Funktion $\psi_1(x)$, welche auf $S(\xi)$ gleich $\psi_2(x)$ genommen wird und die sich in den zu $S(\xi)$ komplementären Intervallen linear ändert, wird unterhalb totalstetig sein in (α, β). Eine stetige Funktion $\varphi_1(x)$, auf $S(\xi)$ gleich $\varphi_2(x)$ und sich linear ändernd in den komplementären Intervallen von $S(\xi)$, ist oberhalb totalstetig in (α, β).

Nach Satz V haben $\psi_1(x)$ und $\varphi_1(x)$ fast überall in (α, β) endliche Ableitungen. Da die Variation von $\psi_1(x)$ in jedem komplementären Intervall von $S(\xi)$ nicht kleiner ist als diejenige von $\varphi_1(x)$, so ist auf der in bezug auf (α, β) komplementären Menge $C(S)$ die Ableitung von ψ_1 größer als oder gleich der Ableitung von φ_1. Auf S muß die Ableitung von ψ_1 (von φ_1) fast überall gleich den unteren Derivierten von ψ_2 (den oberen Derivierten von φ_2) sein. Man sieht nun leicht, daß sich in (α, β) eine Funktion angeben läßt, für die ψ_1 und φ_1 durch Addition bestimmter Konstanten zu adjungierten Ober- und Unterfunktionen werden im Sinne der Definitionen 3 und 4.

Nach Satz VII wird die Differenz $\omega_1(x) \equiv \psi_1(x) - \varphi_1(x)$ nicht-abnehmend sein in (α, β); auch $\omega_2(x)$, welche auf $S(\xi)$ mit $\omega_1(x)$ zusammenfällt, ist nicht-abnehmend in (α, β) und dadurch weiter auch in jedem zu E_2 komplementären abgeschlossenen Teilintervall von (a, b).

E_2 besteht wieder aus einem perfekten Kern K_2 und einem abzählbaren Rest R_2. Wie soeben sieht man, daß K_2 eine auf K_2 nirgends

[22]) Siehe Hobson, loc. cit. [2]), § 81.

dichte abgeschlossene Menge E_3 enthält, in deren in (a, b) liegenden komplementären abgeschlossenen Intervallen $\omega_2(x)$ nicht abnimmt. So fortfahrend kann man K_α definieren für die endlichen und transfiniten Ordnungszahlen (α) der ersten und zweiten Zahlenklasse. Da $K_1 \supset K_2 \ldots K_\omega \supset K_{\omega+1} \ldots$ und diese Mengen auf jeder vorhergehenden derartigen Menge nirgends dicht liegen, so muß eine Zahl β der ersten oder zweiten Zahlenklasse existieren, für die K_β leer ist[23]). Somit ist $\omega_2(x)$ auch nicht-abnehmend in (a, b) [24]).

Definition 15. Zu der in (a, b) definierten Funktion $f(x)$ nimmt man den Wert des *unbestimmten oberen P-Integrals* $\Psi_2(x)$ in einem willkürlichen Punkte x von (a, b) gleich der unteren Schranke aller ψ_2-Werte in diesem Punkte, während das *unbestimmte untere P-Integral* $\Phi_2(x)$ in x definiert wird gleich der oberen Schranke aller φ_2-Werte in x. Wir schreiben

$$\Psi_2(x) \equiv \overline{\int_a^x} (P_2)\, f(x)\, dx = \operatorname{fin\,inf} \psi_2(x)$$

und

$$\Phi_2(x) \equiv \underline{\int_a^x} (P_2)\, f(x)\, dx = \operatorname{fin\,sup} \varphi_2(x).$$

Wenn in (a, b) $\Psi_2(x) = \Phi_2(x)$ ist, so definiert ihr gemeinsamer Wert *das unbestimmte P-Integral* $S_2(x)$ in (a, b).

Satz XV. *Wenn* $S_2(b)$ *existiert* $[= \Psi_2(b) = \Phi_2(b)]$, *so existiert* $S_2(x)$ *für jeden Punkt* x *von* (a, b).

Satz XVI. *Wenn* $S_2(x)$ *in* (a, b) *existiert, so sind für willkürliche* $\psi_2(x)$ *und* $\varphi_2(x)$ *die Differenzen* $\psi_2(x) - S_2(x)$ *und* $S_2(x) - \varphi_2(x)$ *in* (a, b) *nicht-abnehmende Funktionen.*

Satz XVII. *Das unbestimmte P-Integral* $S_2(x)$ *einer Funktion* $f(x)$ *besitzt fast überall eine Ableitung* $= f(x)$.

Die Beweise stimmen völlig überein mit denen der Sätze VIII, IX und X.

Satz D. *Die Definition* 15 *des unbestimmten P-Integrals* $S_2(x)$ *ist äquivalent mit der Definition des Denjoy-Khintchineschen Integrals.*

---

[23]) Siehe Hobson, loc. cit. [2]), § 82.

[24]) Die Gültigkeit des hiermit bewiesenen Satzes hätte es möglich gemacht, *die hier eingeführten Klassen von Ober- und Unterfunktionen anzuwenden in die Definition 5, ohne dadurch die Gültigkeit der Sätze* VIII, IX *und* X *und von Satz* A *zu beeinträchtigen*; diese Klassen würden dann die von Bauer bei beschränkten Funktionen benutzten Klassen umfaßt haben (vgl. § 9). *Der Satz* C *würde nur gelten bleiben unter Hinzufügung der folgenden Bedingung in die Definition 5: „Es existiert mindestens eine Oberfunktion* $\psi_2(x)$, *welche oberhalb totalstetig ist in* (a, b), *und mindestens eine Unterfunktion* $\varphi_2(x)$, *welche unterhalb totalstetig ist in* (a, b).“

Beweis. Wenn $S_2(x)$ für die Funktion $f(x)$ in (a, b) existiert, so muß es eine Folge von Oberfunktionen $\{\psi_2^{(k)}(x)\}$ und eine Folge von Unterfunktionen $\{\varphi_2^{(k)}(x)\}$ geben, derartig, daß $\lim_{k=\infty} \psi_2^{(k)}(b) = \lim_{k=\infty} \varphi_2^{(k)}(b) = S_2(b)$ ist. Aus den Definitionen 13 und 14 folgt, daß (a, b) sich in eine abzählbare Folge von Teilmengen (E_j) zerlegen läßt, derartig, daß gleichzeitig jede der $\{\psi_2^{(k)}\}$ unterhalb totalstetig und jede der $\{\varphi_2^{(k)}\}$ oberhalb totalstetig sind auf einer jeden dieser Teilmengen und dadurch auch auf einer jeden abgeschlossenen Menge M_j, welche aus einer Menge E_j hervorgeht durch Hinzufügung der nicht zu ihr gehörenden Verdichtungspunkte.

Wir definieren nun Funktionen $\{\psi_1^{(kj)}(x)\}$, welche stetig sind für $a \leqq x \leqq b$, die für jede Kombination (kj) auf M_j gleich den zugehörigen $\psi_2^{(k)}$-Werten sind und sich in den zu M_j komplementären Intervallen von (a, b) linear ändern [25]). Auf übereinstimmende Weise leitet man aus den $\{\varphi_2^{(k)}(x)\}$ stetige Funktionen $\{\varphi_1^{(kj)}(x)\}$ ab [25]). Nun läßt sich bei fest gehaltenem j und $k = 1, 2, \ldots$ in (a, b) eine Funktion $g_j(x)$ angeben, bei der die Funktionen $\{\psi_1^{(kj)}(x)\}$ und $\{\varphi_1^{(kj)}(x)\}$ Ober- bzw. Unterfunktionen sind im Sinne der Definitionen 3 und 4 (vgl. den Beweis des Satzes XIV). Daraus folgt, daß $g_j(x)$ ein P-Integral $S_1^{(j)}(x)$ besitzt gemäß Definition 5 [26]), wobei nach Satz C $S_1^{(j)}(x)$ totalstetig sein wird in (a, b). Aus den Sätzen IX und XVI folgt, daß für jeden Punkt ξ von M_j

$$S_1^{(j)}(\xi) = \lim_{k=\infty} \psi_1^{(kj)}(\xi) = \lim_{k=\infty} \psi_2^{(k)}(\xi) = S_2(\xi)$$

ist. Somit wird $S_2(x)$ totalstetig sein auf jedem M_j und dadurch TV in (a, b).

Nun läßt sich die Definition des Denjoy-Khintchineschen Integrals in folgender deskriptiver Form schreiben: „Eine Funktion $f(x)$ hat ein DK-Integral $= G(x)$, wenn: 1. $G(x)$ eine stetige „fonction résoluble" ist, und 2. fast überall eine Ableitung $= f(x)$ hat [27])."

Nach dem Hilfssatze dieses Paragraphen wird eine stetige Funktion, welche TV ist in (a, b), auch eine „fonction résoluble" sein [28]). $S_2(x)$ ist also eine stetige „fonction résoluble" und außerdem hat sie fast überall in (a, b) eine Ableitung $= f(x)$ (Satz XVII). Sie fällt dadurch mit dem unbestimmten DK-Integral $G(x)$ zusammen.

---

[25]) Wir nehmen an, daß in den Endpunkten a und b für jede Kombination (kj) $\psi_1^{(kj)}[\varphi_1^{(kj)}]$ denselben Wert hat wie die zugehörige Funktion $\psi_2^{(k)}[\varphi_2^{(k)}]$.

[26]) Dabei soll die Bedingung der „Beschränktheit" fortbleiben.

[27]) Eine stetige „fonction résoluble" kann definiert werden als eine stetige Funktion mit der Eigenschaft, daß jede perfekte Menge ein Stück enthält, auf dem die Funktion totalstetig sei. — Siehe Khintchine, loc. cit. [3]) und Denjoy, loc. cit. [4]), zweites Zitat, S. 184 (u. 199); erstes Zitat S. 173 u. 175.

[28]) Umgekehrt läßt sich zeigen, daß eine stetige „fonction résoluble" auch TV ist; der Beweis verläuft, wie der des Satzes XIV, mittels transfiniter Induktion.

Wenn umgekehrt $f(x)$ ein DK-Integral $G(x)$ hat in (a, b), so ist $G(x)$ sowohl eine Oberfunktion $\psi_2(x)$ wie eine Unterfunktion $\varphi_2(x)$ [29]. Die Funktion $f(x)$ ist somit auch P-integrierbar im Sinne der Definition 15 und beide Integrale haben denselben Wert.

§ 7.

Es ist nun die Frage, wie man die Klassen der Ober- und Unterfunktionen einschränken soll, damit die Definition des P-Integrals sich identifiziere mit der des speziellen Denjoyschen Integrals. Die Antwort darauf liefern die folgenden Definitionen und der Satz E.

Definition 16. $f(x)$ sei eine im Intervall $(a \leqq x \leqq b)$ definierte Funktion und E eine Menge in diesem Intervall. Wir betrachten endlich viele, nicht übereinandergreifende Intervalle (a_k, b_k), deren Endpunkte zu E gehören sollen, und bilden die Summen

$$\sum_{(k)}\{f(b_k) - f(c_k)\} \quad \text{und} \quad \sum_{(k)}\{f(c_k) - (a_k)\},$$

wobei c_k für jedes fest gewählte k alle möglichen Lagen einnimmt, die den Bedingungen

$$a_k \leqq c_k \leqq b_k$$

genügen. $f(x)$ ist dann *unterhalb totalstetig* auf* E, wenn jene Summen stets einen unteren Limes $\geqq 0$ haben, sobald die Längensumme der Intervalle (a_k, b_k) gegen Null konvergiert.

Definition 17. Ändert man in obiger Definition „unteren Limes $\geqq 0$" in „oberen Limes $\leqq 0$", so entsteht die Bedingung, unter der $f(x)$ *auf* E *oberhalb totalstetig** sein soll.

Definition 17*. Eine auf E unterhalb totalstetige* und oberhalb totalstetige* Funktion wird eine *auf* E *totalstetige** Funktion genannt [30].

Definition 18. $f(x)$ ist *in* (a, b) *unterhalb totalstetig* im verallgemeinerten Sinne* (abgekürzt: UTV*), wenn (a, b) sich in höchstens abzählbar unendlich viele Teilmengen (H_k) zerlegen läßt, auf denen $f(x)$ unterhalb totalstetig* sei.

Definition 19. $f(x)$ ist *in* (a, b) *oberhalb totalstetig* im verallgemeinerten Sinne* (abgekürzt: OTV*), wenn (a, b) sich in höchstens abzählbar unendlich viele Teilmengen (H_k) zerlegen läßt, auf denen $f(x)$ oberhalb totalstetig* sei.

Definition 19*. Eine Funktion, welche UTV* und OTV* ist in (a, b), wollen wir nennen: *totalstetig* auf* (a, b) *im verallgemeinerten Sinne* (abgekürzt: TV*). Eine derartige Funktion hat die charakteristische

[29]) Hier wird die Eigenschaft aus Fußnote [28]) angewandt.

[30]) Vgl. Saks, loc. cit. [13]).

Eigenschaft, daß sich (a, b) zerlegen läßt in eine endliche oder abzählbar unendliche Anzahl von Mengen, auf denen die Funktion totalstetig* ist[30]).

Satz XVIII. *Jede Funktion $f(x)$, welche UTV* ist in (a, b), hat fast überall in (a, b) eine endliche Ableitung und ist somit meßbar.*

Beweis. E sei eine der Mengen, auf denen $f(x)$ unterhalb totalstetig* (abgekürzt: UT*) ist. Durch Hinzufügung der Verdichtungspunkte von E, welche nicht zu E gehörten, entsteht aus E eine abgeschlossene Menge E^*. In einem Punkte ξ von E^*, der auf beiden Seiten Grenzpunkt von E ist, betrachten wir den unteren und den oberen Limes von $f(x)$ in bezug auf (a, b); in einem Punkte ξ_1 von E^*, der nur linker [oder rechter] Grenzpunkt von E ist, betrachten wir den unteren und den oberen Limes von $f(x)$ in bezug auf die Punkte von (a, b), welche $\leqq \xi_1$ [bzw. $\geqq \xi_1$] sind im Falle, daß ξ_1 zu E gehört, und $< \xi_1$ [bzw. $> \xi_1$] sind im Falle, daß ξ_1 zu E nicht gehört; in einem Punkte ξ_2 von E^*, welcher isolierter Punkt von E ist, definieren wir den gemeinsamen Wert der unteren und oberen Limes von $f(x)$ gleich $f(\xi_2)$. Die auf diese Weise auf E^* definierten unteren und oberen Limesfunktionen können nur in endlich vielen Punkten von E^* unendlich sein (vgl. den Anfang des Beweises von Satz XI)[31]). Es sei H die Menge dieser Punkte.

Betrachten wir die Menge E_1^* aller zu $E^* - H$ gehörenden ξ-Punkte[32]); die Differenzmenge von $E^* - H$ und E_1^* ist abzählbar. Für jeden Punkt von E_1^* muß eine Umgebung $\mathfrak{U}$ existieren, derartig, daß $f(x)$ in dieser Umgebung beschränkt ist. Die Menge E_1^* läßt sich, nach dem Satze von Lindelöf[17]), von einer abzählbaren Folge von derartigen Umgebungen $(\mathfrak{U}_k)$ überdecken. Auf jeder Menge $\mathfrak{U}_k \cdot E_1^*$ wird $f(x)$ unterhalb totalstetig* sein[33]).

---

[31]) Diejenigen Punkte von E^*, in denen diese unteren und oberen Limesfunktionen nicht beide endlich sind, können nicht ξ-Punkte sein; in den ξ_1-Punkten dieser Art, welche linke Grenzpunkte sind, haben beide Limesfunktionen den Wert $+\infty$, und in denjenigen, welche rechte Grenzpunkte sind, haben sie den Wert $-\infty$. Vgl. Fußn. [16]).

[32]) D. h. also Punkte, die auf beiden Seiten Grenzpunkt von E sind. Aus Fußn. [31]) geht hervor, daß H keine ξ-Punkte enthalten kann, so daß $E^* - H$ alle ξ-Punkte enthält.

[33]) Wenn ξ' und ξ'' Punkte von $\mathfrak{U}_k \cdot E_1^*$ sind und p ein willkürlicher innerer Punkt des Intervalls (ξ', ξ'') ist, so lassen sich immer Punkte α, β, γ und δ von E angeben, welche den Ungleichungen

$$\alpha < \xi' < \beta < p < \gamma < \xi'' < \delta$$

genügen und wobei die Differenzen $\beta - \alpha$ und $\delta - \gamma$ beliebig klein sind; es läßt sich schreiben:

$$f(p) - f(\xi') = \{f(p) - f(\beta)\} + \{f(\beta) - f(\xi')\}$$

und

$$f(\xi'') - f(p) = \{f(\xi'') - f(\gamma)\} + \{f(\gamma) - f(p)\}.$$

Mit Hilfe von derartigen Identitäten für alle möglichen Intervalle (ξ', ξ'') und alle möglichen Punkte p in diesen Intervallen läßt sich leicht die Behauptung des Textes beweisen.

Es sei H_k die Teilmenge $\mathfrak{U}_k E_1^*$, in deren Punkten die rechte untere Derivierte von $f(x)$ $[D_+ f(x)]$ gleich $-\infty$ ist. Nehmen wir an, daß ihr äußeres Maß positiv sei. Auch auf H_k wird $f(x)$ unterhalb totalstetig* sein. Bei willkürlich positivem ε existiert somit ein positives δ, derartig, daß für endlich viele Intervalle (a_j, b_j), welche nicht übereinandergreifen und deren Endpunkte zu H_k gehören sollen, immer, wie auch für jedes j der Punkt c_j im abgeschlossenen Intervall (a_j, b_j) gewählt werden möge,

$$\sum_{(j)} \{f(c_j) - f(a_j)\} > -\varepsilon \tag{9a}$$

ist, wenn nur

$$\sum_{(j)} m(a_j, b_j) < \delta \tag{9b}$$

bleibt.

Da H_k in fast jedem ihrer Punkte eine äußere Dichte $= 1$ hat[34]), existiert ein Intervall (α, β), dessen Endpunkte zu H_k gehören, das eine Teilmenge T_k von H_k enthält von positivem äußeren Maße und dessen Länge $< \delta$ ist.

Zu jedem Punkte x von T_k, ausgenommen den Punkt β, gehört eine abzählbare Folge von Intervallen (x, x_n), welche in (α, β) liegen, sich mit zunehmendem n in x zusammenziehen und für die

$$f(x_n) - f(x) < -\frac{A}{m_a(T_k)} \cdot (x_n - x)$$

ist, wobei A eine willkürliche positive Zahl darstellt. Wenn $0 < p < 1$ ist, so läßt sich nach dem Überdeckungssatze von Vitali[7]) eine endliche Anzahl von nicht übereinandergreifenden Intervallen aus der Menge aller (x, x_n) wählen, deren Gesamtmaß größer als $(1-p) \cdot m_a(T_k^*)$ ist und wobei dann die Summe der Funktionsdifferenzen in den Endpunkten dieser Intervalle

$$< -\frac{A}{m_a(T_k)} \cdot (1-p)\, m_a(T_k) \quad \text{oder} \quad < -(1-p) \cdot A$$

sein wird. Die linken Endpunkte dieser Intervalle teilen (α, β) in eine endliche Anzahl von aneinandergrenzenden Intervallen, deren Endpunkte zu T_k gehören; nun hätte man vorher, bei fest gewählten p und ε, die Zahl A so bestimmen können, daß

$$(1-p) \cdot A > \varepsilon$$

---

[34]) Siehe Sierpiński, Fund. Math. 4 (1923), S. 125. Eine lineare Menge E hat im Punkte ξ die äußere Dichte 1, wenn für jede Folge von ξ im Innern oder auf dem Rande enthaltenden und sich in ξ zusammenziehenden Intervallen (J_k)

$$\lim_{k=\infty} \frac{m_a(E \cdot J_k)}{m(J_k)} = 1$$

ist.

wäre. Dies führt zu einem Widerspruch mit (9a) und (9b). Somit kann $D_+ f$ nur gleich $-\infty$ sein auf einer Nullmenge von $\mathfrak{U}_k \cdot E_1^*$.

Auf gleiche Weise läßt sich zeigen, daß auch $D_- f = -\infty$ auf E_1^* nur in den Punkten einer Nullmenge möglich ist.

Anwendung des Satzes von Denjoy-Saks[8]) auf $f(x)$ im offenen Intervall $\mathfrak{U}_k$ liefert schließlich, daß $f(x)$ in fast allen Punkten von $\mathfrak{U}_k \cdot E_1^*$ eine endliche Ableitung haben muß. Somit muß eine endliche Ableitung auch existieren fast überall auf E_1^*, auf E^* und schließlich in (a, b).

Satz XIX. *Wenn zwei endliche Funktionen $f(x)$ und $g(x)$ UTV* bzw. OTV* sind in (a, b), so wird die Differenz $f(x) - g(x)$ UTV* sein in (a, b).*

§ 8.

Definition 20. $f(x)$ sei eine in (a, b) definierte Funktion. Eine *zu $f(x)$ im Intervall (a, b) adjungierte Oberfunktion* $\psi_2^*(x)$ soll stetig sein für $a \leqq x \leqq b$ und daneben den folgenden Bedingungen genügen: a) $\psi_2^*(a) = 0$; b) $\psi_2^*(x)$ ist UTV* in (a, b); c) die beiden unteren Derivierten von $\psi_2^*(x)$ sind fast überall in (a, b) $\geqq f(x)$.

Definition 21. Eine *zu $f(x)$ in (a, b) adjungierte Unterfunktion* $\varphi_2^*(x)$ ist stetig für $a \leqq x \leqq b$ und genügt den weiteren Bedingungen: a) $\varphi_2^*(a) = 0$; b) $\varphi_2^*(x)$ ist OTV* in (a, b); c) die beiden oberen Derivierten von $\varphi_2^*(x)$ sind fast überall in (a, b) $\leqq f(x)$.

Satz XX. *Die Differenz einer jeden Oberfunktion $\psi_2^*(x)$ und einer jeden Unterfunktion $\varphi_2^*(x)$, adjungiert zu der in (a, b) definierten Funktion $f(x)$, ist nicht-abnehmend in (a, b).*

Dies ist ein Spezialfall des Satzes XIV.

Definition 22. Zu der in (a, b) definierten Funktion $f(x)$ nimmt man den Wert des *unbestimmten oberen P-Integrals* $\Psi_2^*(x)$ in einem willkürlichen Punkte x von (a, b) gleich der unteren Schranke aller ψ_2^*-Werte in diesem Punkte, während das *unbestimmte untere P-Integral* $\Phi_2^*(x)$ in x definiert wird gleich der oberen Schranke aller φ_2^*-Werte in x. Wir schreiben

$$\Psi_2^*(x) \equiv \overline{\int_a^x} (P_2^*) f(x)\, dx = \operatorname{fin\,inf} \psi_2^*(x)$$

und

$$\Phi_2^*(x) \equiv \underline{\int_a^x} (P_2^*) f(x)\, dx = \operatorname{fin\,inf} \varphi_2^*(x).$$

Wenn in (a, b) $\Psi_2^*(x) = \Phi_2^*(x)$ ist, so definiert ihr gemeinsamer Wert *das unbestimmte P-Integral* $S_2^*(x)$ in (a, b).

Satz XXI. *Wenn* $S_2^*(b)$ *existiert* $[= \Psi_2^*(b) = \Phi_2^*(b)]$, *so existiert* $S_2^*(x)$ *in jedem Punkte* x *von* (a, b).

Satz XXII. *Wenn* $S_2^*(x)$ *in* (a, b) *existiert, so sind, für willkürliche* $\psi_2^*(x)$ *und* $\varphi_2^*(x)$, *die Differenzen* $\psi_2^*(x) - S_2^*(x)$ *und* $S_2^*(x) - \varphi_2^*(x)$ *in* (a, b) *nicht-abnehmende Funktionen.*

Satz XXIII. *Das unbestimmte P-Integral* $S_2^*(x)$ *einer Funktion* $f(x)$ *besitzt fast überall eine Ableitung* $= f(x)$.

Satz E. *Die Definition* 22 *des unbestimmten P-Integrals* $S_2^*(x)$ *ist äquivalent mit der Definition des speziellen Denjoyschen Integrals.*

Beweis. Wenn $S_2^*(x)$ für die Funktion $f(x)$ in (a, b) existiert, so muß es eine Folge von Oberfunktionen $\{\psi_2^{*(k)}(x)\}$ und eine Folge von Unterfunktionen $\{\varphi_2^{*(k)}(x)\}$ geben, derartig, daß $\lim_{k=\infty} \psi_2^{*(k)}(b) = \lim_{k=\infty} \varphi_2^{*(k)}(b) = S_2^*(b)$ ist. Aus den Definitionen 20 und 21 folgt, daß (a, b) sich in eine abzählbare Folge von Teilmengen (E_j) zerlegen läßt, derartig, daß gleichzeitig jede der $\{\psi_2^{*(k)}\}$ unterhalb totalstetig* und jede der $\{\varphi_2^{*(k)}\}$ oberhalb totalstetig* sind auf einer jeden dieser Teilmengen.

Nun muß $S_2^*(x)$ totalstetig* sein auf jedem E_j. Sonst existierte für eine dieser Mengen, $E_{j'}$, eine z. B. positive Zahl Δ und eine Folge von Mengen (M_n), welche jede für sich von endlich vielen, nicht übereinandergreifenden, abgeschlossenen Intervallen $(a_p^{(n)}, a_{p+1}^{(n)})$ gebildet würden, deren Endpunkte zu $E_{j'}$ gehörten und welche Punkte von $x_p^{(n)}$ enthielten, so daß gleichzeitig

$$(10) \qquad \lim_{n=\infty} \sum_{(p)} [S_2^*(x_p^{(n)}) - S_2^*(a_p^{(n)})] = \Delta \qquad \text{und} \qquad \lim_{n=\infty} m(M_n) = 0$$

wäre.

$\varphi_2^{*(k')}(x)$ sei eine adjungierte Unterfunktion aus der Folge $\{\varphi_2^{*(k)}(x)\}$ mit

$$S_2^*(b) - \varphi_2^{*(k')}(b) < \frac{1}{2}\Delta .$$

Nach Satz XXII ist $S_2^*(x) - \varphi_2^{*(k')}(x)$ eine nicht-abnehmende Funktion. Daraus folgt bei willkürlich, aber fest gewähltem n:

$$(11) \qquad \sum_{(p)} [S_2^*(x_p^{(n)}) - S_2^*(a_p^{(n)})] - \sum_{(p)} [\varphi_2^{*(k')}(x_p^{(n)}) - \varphi_2^{*(k')}(a_p^{(n)})]$$
$$\leqq S_2^*(b) - \varphi_2^{*(k')}(b) < \frac{1}{2}\Delta .$$

Aus (10) und (11) folgt

$$\limsup_{n=\infty} \sum_{(p)} [\varphi_2^{*(k')}(x_p^{(n)}) - \varphi_2^{*(k')}(a_p^{(n)})] > \frac{1}{2}\Delta .$$

Dies ist jedoch unmöglich, denn $\varphi_2^{*(k')}$ ist oberhalb totalstetig* auf $E_{j'}$.

Auch die Annahme eines negativen Δ würde zu einem Widerspruch führen. $S_2^*(x)$ ist somit totalstetig* auf jedem E_j, und dadurch TV* in (a, b).

Der Definition des speziellen Denjoyschen Integrals kann man eine deskriptive Form geben: „Eine Funktion $f(x)$ hat ein spezielles Denjoysches Integral $= H(x)$, wenn 1. $H(x)$ eine stetige ‚fonction complètement résoluble' ist und 2. fast überall eine Ableitung $= f(x)$ hat[35])."

Nach dem Hilfssatze des § 6 wird eine stetige Funktion, welche TV^* ist in (a, b), auch „complètement résoluble" sein[36]). $S_2^*(x)$ ist also eine stetige „fonction complètement résoluble", und außerdem hat sie fast überall eine Ableitung $= f(x)$ (Satz XXIII). Sie fällt dadurch mit dem unbestimmten speziellen D-Integral $H(x)$ zusammen.

Wenn umgekehrt $f(x)$ ein spezielles D-Integral $H(x)$ hat, so ist $H(x)$ sowohl eine Oberfunktion $\psi_2^*(x)$ wie eine Unterfunktion $\varphi_2^*(x)$.[37]) Die Funktion $f(x)$ ist somit auch P-integrierbar im Sinne der Definition 22, und beide Integrale haben denselben Wert.

§ 9.

Die für $a \leqq x \leqq b$ zu $f(x)$ gehörenden *Perron-Bauerschen Oberfunktionen* sind 1. stetig; 2. $= 0$ in a; 3. haben ihre unteren Derivierten $\geqq f(x)$,[38]) während 4. diese unteren Derivierten $\neq -\infty$ sind. Aus 1. und 2. läßt sich folgern, *daß sie in* (a, b) UTV^* *sind.*

Beweis. Es sei $G(x)$ eine derartige Oberfunktion. Da die unteren Derivierten $\neq -\infty$ sind, wird, nach einem Satze von Denjoy[39]), jede perfekte Menge E ein Stück $\tilde{\omega}$ (mit den Endpunkten c und d) enthalten, derartig, daß für jedes in (c, d) liegende Intervall (ξ, η), dessen linker oder rechter Endpunkt (oder beide) zu $\tilde{\omega}$ gehört,

$$\frac{G(\eta) - G(\xi)}{\eta - \xi}$$

---

[35]) Eine stetige „fonction complètement résoluble" kann definiert werden als eine stetige Funktion mit der Eigenschaft, daß jede perfekte Menge ein Stück enthält, auf dem die Funktion totalstetig* sei, oder, was dasselbe bedeutet, auf dem die Funktion totalstetig sei, während die Oszillationen der Funktion in den komplementären Intervallen des Stückes eine konvergente Reihe bilden. — Siehe Denjoy, loc. cit. [2]) und loc. cit. [4]), zweites Zitat, S. 184 u. 199, 200; erstes Zitat, S. 173 u. 175; oder unsere im Nieuw Archief (Amsterdam) erscheinende Arbeit: Sur deux classes de fonctions, introduites par M. Denjoy, th. 13, 15, 10.

[36]) Umgekehrt ist eine stetige „fonction complètement résoluble" auch TV^*.

[37]) Hier wird die Eigenschaft der vorigen Fußnote angewandt.

[38]) Die Hinzufügung „fast überall" beeinträchtigt zwar den Umfang der Klasse der Oberfunktionen, jedoch nicht den Umfang der auf ihnen und auf den übereinstimmenden Unterfunktionen gegründeten P.-B. Integraldefinition. Siehe Kamke, loc. cit. [1]), S. 132–135.

[39]) Siehe Lebesgue, loc. cit. [13]), S. 220.

oberhalb einer festen Schranke bleibt[40]). Daraus folgt unmittelbar, daß $G(x)$ unterhalb totalstetig* ist auf $\bar{\omega}$.

Nun verläuft der Beweis weiter mittels transfiniter Induktion. Die Punkte von (a, b), in deren Umgebung $G(x)$ nicht unterhalb totalstetig* ist, bilden eine nirgends dichte, abgeschlossene Menge E_1. Die zu E_1 komplementären Intervalle von (a, b) enthalten abzählbar unendlich viele Mengen, auf denen $G(x)$ UT* ist. E_1 ist die Summe einer abzählbaren Menge M_1 und eines perfekten Kerns K_1.[32]) Die einzelnen Punkte von M_1 sind zu betrachten als Mengen, auf denen $G(x)$ UT* ist. Mit K_1 verfahren wir wie mit (a, b) und erhalten so eine perfekte Menge K_2. So fortfahrend, haben wir eine Folge von perfekten Mengen $K_1 \supset K_2 \supset \ldots \supset K_n \ldots$, welche für ein bestimmtes n verschwinden oder eine abgeschlossene Menge E_ω gemeinsam haben[41]), welche zerlegt wird in die abzählbare Menge M_ω und den perfekten Kern K_ω. Darauf wird $K_{\omega+1}$ gebildet. Da die perfekten Mengen auf jeder vorhergehenden nirgends dicht liegen, muß es eine erste Ordnungszahl Ω der zweiten Zahlenklasse geben, für die K_Ω leer ist[33]). Die verschiedenen, in diesem Prozesse gebildeten Mengen, auf denen $G(x)$ UT* ist, und die einzelnen Punkte der Mengen (M_n) liefern abzählbar unendlich viele Mengen, auf denen $G(x)$ UT* ist und welche zusammen (a, b) aufbauen. Somit ist $G(x)$ in (a, b) UTV*.

Die *P.-B. Unterfunktionen* von $f(x)$ sind: 1. stetig; 2. $= 0$ in a; 3. haben in (a, b) ihre oberen Derivierten $\leqq f(x)$, während: 4. die oberen Derivierten $\neq +\infty$ sind. Aus 1. und 4. folgt, *daß sie in* (a, b) *OTV** *sind*[42]).

Da also die P.-B. Ober- und Unterfunktionen in den Klassen der $\{\psi_2^*(x)\}$ und der $\{\varphi_2^*(x)\}$ enthalten sind, gibt Anwendung des Satzes E unmittelbar:

Jede im Perron-Bauerschen Sinne integrierbare Funktion hat auch ein spezielles Denjoysches Integral.

Ein von Looman[12]) und Alexandroff[12]), unabhängig voneinander, erreichtes Resultat ist hiermit aufs neue, zwar auf indirekte Weise, bewiesen[43]).

---

[40]) Dies gilt auch, wenn in 4. angenommen wird, daß die unteren Derivierten nur $-\infty$ werden können auf einer abzählbaren Menge (siehe Lebesgue, loc. cit. [13]), S. 225). Diese Annahme wird auch den weiteren Teil des Beweises ungeändert lassen; derartig verallgemeinerte Oberfunktionen sind also ebenfalls UTV* in (a, b).

[41]) Siehe Hobson, loc. cit. [2]), § 67.

[42]) Dies gilt auch, wenn in 4. angenommen wird, daß die oberen Derivierten nur $+\infty$ werden können auf einer abzählbaren Menge.

[43]) Beweise für die Umkehrung dieses Satzes gaben Hake (loc. cit. [11]), §§ 2–4) und Alexandroff (loc. cit. [12])). — Wenn man die in den Fußnoten [40]) und [42]) angegebenen Verallgemeinerungen der Bedingungen 4. in die Definitionen der P.-B. Ober-

(Fortsetzung der Fußnote [43]) auf nächster Seite.)

Viertes Kapitel.

§ 10.

Da die in § 6 eingeführten Klassen der $\{\psi_3(x)\}$ und der $\{\varphi_2(x)\}$ nicht die Möglichkeit gaben, das P-Integral mit dem allgemeinen D-Integral zusammenfallen zu lassen, wird es notwendig sein, die Klassen noch einmal zu erweitern.

Definition 23. $f(x)$ sei eine in (a, b) definierte Funktion. Eine *zu* $f(x)$ *im Intervall* (a, b) *adjungierte Oberfunktion* $\psi_3(x)$ soll stetig sein für $a \leqq x \leqq b$ und daneben den folgenden Bedingungen genügen: a) $\psi_3(a) = 0$; b) $\psi_3(x)$ ist UTV in (a, b); c) es gibt eine Zerlegung von (a, b) in höchstens abzählbar unendlich viele Teilmengen (H_k), derartig, daß $\psi_3(x)$ auf jedem H_k fast überall untere Derivierten „in bezug auf dieses H“ hat $\geqq f(x)$.[44])

Definition 24. Eine *zu* $f(x)$ *in* (a, b) *adjungierte Unterfunktion* $\varphi_3(x)$ ist stetig für $a \leqq x \leqq b$ und genügt den weiteren Bedingungen: a) $\varphi_3(a) = 0$; b) $\varphi_3(x)$ ist OTV in (a, b); c) es gibt eine Zerlegung von (a, b) in höchstens abzählbar unendlich viele Teilmengen (H_k), derartig, daß $\varphi_3(x)$ auf jedem H_k fast überall obere Derivierten „in bezug auf dieses H_k“ hat $\leqq f(x)$.[44])

Aus Satz XI folgt, daß jedes $\psi_3(x)$ und jedes $\varphi_3(x)$ fast überall endliche approximative Ableitungen haben. Ohne dieses Resultat dabei anzuwenden, läßt sich zeigen:

Satz XXIV. *Wenn eine Funktion* $f(x)$ *in* (a, b) *gleichzeitig eine Oberfunktion* $\psi_3(x)$ *und eine Unterfunktion* $\varphi_3(x)$ *hat, so gibt es einen maßgleichen Kern* K *von* (a, b), *in dessen Punkten endliche approximative Ableitungen,* $D_{\text{appr.}}\psi_3(x)$ *bzw.* $D_{\text{appr.}}\varphi_3(x)$ *existieren, welche in jedem dieser Punkte der Ungleichung*

$$D_{\text{appr.}}\psi_3(x) \geqq D_{\text{appr.}}\varphi_3(x) \tag{12}$$

genügen.

---

und Unterfunktionen zuläßt und in die Bedingungen 3. „fast überall“ hinzufügt, so erhält man einen Integralbegriff, der sich genau mit dem ursprünglichen P.-B. Integralbegriff deckt. (Das wurde übrigens schon von Looman, loc. cit. [12]), S. 153 bemerkt.) Es folgt daraus unmittelbar, *daß eine Ableitung* $f'(x)$, *welche in* (a, b) *endlich ist, ausgenommen in den Punkten einer höchstens abzählbaren Menge, ein P.-B. Integral hat gleich* $f(x) - f(a)$.

[44]) Wir nehmen unter c) an, daß auch aus einem einzelnen Punkte oder aus endlich vielen Punkten aufgebaute Mengen zugelassen sind; da dann die unteren und oberen Derivierten in bezug auf die Menge nirgends existieren, ist die geforderte Bedingung als erfüllt zu betrachten.

Beweis. Aus Definition 23, c) und 24, c) folgt, daß (a, b) sich in abzählbar unendlich viele Mengen (M_k) zerlegen läßt, derartig, daß gleichzeitig in fast jedem Punkte eines M_k die in bezug auf dieses M_k genommenen unteren Derivierten von $\psi_3(x) \geqq f(x)$ und die in bezug auf das M_k genommenen oberen Derivierten von $\varphi_3(x) \leqq f(x)$ sind. Wenn man ψ_3 und φ_3 nur für die Punkte von M_k betrachtet, so liefert Anwendung des Satzes von Denjoy-Saks[8]), daß fast überall auf M_k ψ_3 und φ_3 in bezug auf diese Menge endliche Ableitungen, $D_{M_k}\psi_3(x)$ und $D_{M_k}\varphi_3(x)$ haben, die der Ungleichung

$$D_{M_k}\psi_3(x) \geqq D_{M_k}\varphi_3(x) \tag{13}$$

genügen. Da ψ_3 und φ_3 stetig sind, bleibt (13) gültig in fast jedem Punkte von M_k, wenn man die Ableitungen definiert in bezug auf die Punkte der abgeschlossenen Menge M_k^*, welche aus M_k hervorgeht durch Hinzufügung seiner Verdichtungspunkte.

M_k^* ist meßbar und hat somit in fast jedem seiner Punkte die Dichte 1. Aus dem Vorhergehenden folgt dadurch (12) für fast jeden Punkt von M_k, also auch für fast jeden Punkt von (a, b), wobei die approximativen Ableitungen in einem derartigen Punkte außerdem endlich sind[45]).

Satz XXV. *Die Differenz einer jeden Oberfunktion $\psi_3(x)$ und einer jeden Unterfunktion $\varphi_3(x)$, adjungiert zu der in (a, b) definierten Funktion $f(x)$, ist nicht-abnehmend.*

Beweis. Dieser stimmt überein mit dem des Satzes XIV.

Nach dem Hilfssatze des § 6 enthält (a, b) eine nirgends dichte, abgeschlossene Menge E_1 mit der Eigenschaft, daß jeder Punkt ξ der in bezug auf (a, b) komplementären Menge $C(E_1)$ eine Umgebung $\mathfrak{U}(\xi)$ hat, in der $\psi_3(x)$ und $\varphi_3(x)$ unterhalb bzw. oberhalb totalstetig sind auf einer in $\mathfrak{U}(\xi)$ überall dicht liegenden Menge. Aus der Stetigkeit von ψ_3 und φ_3 folgt dann, daß sie auch auf $\mathfrak{U}(\xi)$ unterhalb bzw. oberhalb totalstetig sind. Nach Satz V haben ψ_3 und φ_3 fast überall in $\mathfrak{U}(\xi)$ endliche Ableitungen $\psi_3'(x)$ und $\varphi_3'(x)$, und aus Satz XXIV folgt dadurch, daß fast überall in $\mathfrak{U}(\xi)$:

$$\psi_3'(x) \geqq \varphi_3'(x)$$

sein muß.

Nach Satz VII wird $\omega_3(x) \equiv \psi_3(x) - \varphi_3(x)$ nicht-abnehmend sein in $\mathfrak{U}(\xi)$ und in jedem zu E_1 komplementären, abgeschlossenen Intervall von (a, b).

E_1 besteht aus dem perfekten Kerne K_1 und einem abzählbaren Rest R_1; $\omega_3(x)$ ist nicht-abnehmend für jedes zu K_1 komplementäre, abgeschlossene Intervall von (a, b).

[45]) Es zeigt sich daneben, daß eine Funktion, welche ein $\psi_3(x)$ und ein $\varphi_3(x)$ hat, fast überall endlich sein muß.

K_1 enthält (Hilfssatz von § 6) eine auf K_1 nirgends dichte, abgeschlossene Menge E_2 mit der Eigenschaft, daß zu jedem Punkte ξ von $K_1 - E_2$ ein Stück $S(\xi)$ von K_1 existiert, dessen Endpunkte α und β zugleich Endpunkte eines ξ im Innern enthaltenden Intervalls sind und auf dem $\psi_3(x)$ unterhalb und $\varphi_3(x)$ oberhalb totalstetig sind. Eine stetige Funktion $\psi_1(x)$, welche auf $S(\xi)$ gleich $\psi_3(x)$ genommen wird und die sich in den zu $S(\xi)$ komplementären Intervallen linear ändert, wird unterhalb totalstetig sein in (α, β). Eine stetige Funktion $\varphi_1(x)$, auf $S(\xi)$ gleich $\varphi_3(x)$ und sich linear ändernd in den komplementären Intervallen von $S(\xi)$, ist oberhalb totalstetig in (α, β).

Nach Satz V haben $\psi_1(x)$ und $\varphi_1(x)$ fast überall in (α, β) endliche Ableitungen, und auf der in bezug auf (α, β) komplementären Menge $C(S)$ ist

$$\psi_1'(x) \geqq \varphi_1'(x).$$

Auf S muß, nach Satz XXIV, fast überall

$$\psi_1'(x) = D_{\text{appr.}}\,\psi_3(x) \geqq D_{\text{appr.}}\,\varphi_3(x) = \varphi_1'(x)$$

sein. In (α, β) läßt sich eine Funktion angeben, für die ψ_1 und φ_1 durch Addition bestimmter Konstanten Ober- und Unterfunktionen werden im Sinne der Definitionen 3 und 4.

Nach Satz VII wird $\omega_1(x) \equiv \psi_1(x) - \varphi_1(x)$ nicht-abnehmend sein in (α, β). Dasselbe gilt also auch für $\omega_3(x)$ in (α, β) und dadurch weiter auch in jedem zu E_2 komplementären, abgeschlossenen Teilintervall von (a, b).

Das Beweisverfahren transfinit fortsetzend, findet man schließlich, daß $\omega_3(x)$ auch in (a, b) nicht abnimmt[46]).

Definition 25. Zu der in (a, b) definierten Funktion $f(x)$ nimmt man den Wert des *unbestimmten oberen P-Integrals* $\Psi_3(x)$ im willkürlichen Punkte x von (a, b) gleich der unteren Schranke aller ψ_3-Werte in x und den Wert des *unbestimmten unteren P-Integrals* $\Phi_3(x)$ gleich der oberen Schranke aller φ_3-Werte in x. Wir schreiben:

$$\Psi_3(x) \equiv \overline{\int_a^x} (P_3) f(x)\,dx = \text{fin inf}\,\psi_3(x)$$

und

$$\Phi_3(x) \equiv \underline{\int_a^x} (P_3) f(x)\,dx = \text{fin sup}\,\varphi_3(x).$$

Wenn in (a, b) $\Psi_3(x) = \Phi_3(x)$ ist, so definiert ihr gemeinsamer Wert *das unbestimmte P-Integral* $S_3(x)$ in (a, b).

Satz XXVI. *Wenn* $S_3(b)$ *existiert* $[= \Psi_3(b) = \Phi_3(b)]$, *so existiert* $S_3(x)$ *für jeden Punkt* x *von* (a, b).

---

[46]) *Hier läßt sich für die Sätze* A *und* C *eine mit Fußnote* [24]) *übereinstimmende Bemerkung machen.*

Satz XXVII. *Wenn $S_3(x)$ in (a, b) existiert, so sind, für willkürliche $\psi_3(x)$ und $\varphi_3(x)$ die Differenzen: $\psi_3(x) - S_3(x)$ und $S_3(x) - \varphi_3(x)$ in (a, b) nicht-abnehmende Funktionen.*

Satz XXVIII. *Das unbestimmte P-Integral $S_3(x)$ einer Funktion $f(x)$ besitzt fast überall eine approximative Ableitung $= f(x)$.*

Beweis. Es gibt eine Folge von Oberfunktionen $\{\psi_3^{(k)}(x)\}$ und eine Folge von Unterfunktionen $\{\varphi_3^{(k)}(x)\}$, derartig, daß

$$\lim_{k=\infty} \psi_3^{(k)}(b) = \lim_{k=\infty} \varphi_3^{(k)}(b) = S_3(b)$$

ist. Aus den Definitionen 23 und 24 folgt, daß (a, b) sich in eine abzählbare Folge von Teilmengen (E_j) zerlegen läßt, derartig, daß gleichzeitig jede der $\{\psi_3^{(k)}\}$ unterhalb totalstetig und jede der $\{\varphi_3^{(k)}\}$ oberhalb totalstetig sind auf einer jeden abgeschlossenen Menge M_j, welche aus einer Menge E_j hervorgeht durch Hinzufügung der nicht zu ihr gehörenden Verdichtungspunkte.

Funktionen $\{\psi_1^{(kj)}(x)\}$, welche stetig sind für $a \leqq x \leqq b$, die auf der zugehörigen Menge M_j gleich der zugehörigen Funktion $\psi_3^{(k)}(x)$ sind und die sich in den zu dieser Menge komplementären Intervallen von (a, b) linear ändern[47]), sind unterhalb totalstetig in (a, b). Auf übereinstimmende Weise leitet man aus den $\{\varphi_3^{(k)}(x)\}$ in (a, b) stetige und oberhalb totalstetige Funktionen $\{\varphi_1^{(kj)}(x)\}$ ab[47]). Nun läßt sich bei festgehaltenem j und $k = 1, 2, \ldots$ in (a, b) eine Funktion $g_j(x)$ angeben, welche auf M_j mit $f(x)$ zusammenfällt und bei der die Funktionen $\{\psi_1^{(kj)}(x)\}$ und $\{\varphi_1^{(kj)}(x)\}$ Ober- bzw. Unterfunktionen sind im Sinne der Definitionen 3 und 4 (vgl. den Beweis des Satzes XXV). $g_j(x)$ besitzt ein P-Integral $S_1^{(j)}(x)$ gemäß Definition 5[48]), und nach Satz X (s. § 4) wird fast überall in (a, b):

$$D S_1^{(j)}(x) = g_j(x),$$

also auch fast überall auf M_j:

$$D S_1^{(j)}(x) = f(x) \tag{14}$$

sein. Aus den Sätzen IX und XXVII folgt für jeden Punkt x von M_j:

$$S_1^{(j)}(x) = \lim_{k=\infty} \psi_1^{(kj)}(x) = \lim_{k=\infty} \psi_3^{(k)}(x) = S_3(x). \tag{15}$$

Da M_j in fast jedem ihrer Punkte eine Dichte 1 hat, folgt aus (14) und (15), daß $S_3(x)$ fast überall auf M_j eine approximative Ableitung $= f(x)$ hat. Dasselbe gilt also auch in (a, b).

Satz F. *Die Definition des P-Integrals $S_3(x)$ ist äquivalent mit der Definition des allgemeinen Denjoyschen Integrals.*

---

[47]) Wir nehmen an, daß in den Endpunkten a und b für jede Kombination (kj) $\psi_1^{(kj)}$ [$\varphi_1^{(kj)}$] denselben Wert hat wie die zugehörige Funktion $\psi_3^{(k)}$ [$\varphi_3^{(k)}$].

[48]) Dabei soll die Bedingung der „Beschränktheit" fortbleiben.

Beweis. $S_3(x)$ existiere in (a, b) für die Funktion $f(x)$. Aus dem Beweise des vorigen Satzes geht hervor, daß $S_3(x)$ TV ist in (a, b) (vgl. auch den Beweis von Satz D).

Die Definition des allgemeinen D-Integrals hat die folgende, deskriptive Form: „Eine Funktion $f(x)$ hat ein allgemeines D-Integral $= J(x)$, wenn 1. $J(x)$ eine stetige ‚fonction résoluble' ist, und 2. fast überall eine approximative Ableitung $= f(x)$ hat" [49]).

Nach dem Hilfssatze von § 6 ist $S_3(x)$ eine „fonction résoluble". Hieraus und aus Satz XXVIII folgt, daß $S_3(x)$ mit $J(x)$ zusammenfällt.

Nehmen wir, umgekehrt, an, daß $f(x)$ in (a, b) ein allgemeines D-Integral $J(x)$ habe. $J(x)$ ist eine stetige „fonction résoluble"; daraus folgt, daß sie TV ist in (a, b).[28]) Es existiert somit eine Zerlegung von (a, b) in eine abzählbare Folge von Mengen (M_j), auf denen $J(x)$ totalstetig ist. Aus ihrer Stetigkeit folgt, daß sie auch totalstetig sein wird auf jeder Menge M_j^*, welche aus den Punkten des zugehörigen M_j nebst allen seinen Verdichtungspunkten aufgebaut ist. Da sich wieder eine in (a, b) totalstetige Funktion bilden läßt, welche auf einer willkürlichen Menge M_j^* mit $J(x)$ zusammenfällt, folgt aus Satz V, daß $J(x)$ auf jeder Menge M_j^* fast überall endliche Ableitungen hat in bezug auf M_j^*. Da M_j^* in fast jedem ihrer Punkte eine Dichte 1 und $J(x)$ fast überall auf M_j^* eine approximative Ableitung $= f(x)$ hat, so wird die Ableitung von $J(x)$ in bezug auf M_j^* auch fast überall auf dieser Menge $= f(x)$ sein; dasselbe gilt also fast überall auf jeder Menge M_j für die Ableitungen in bezug auf M_j. Dadurch genügt $J(x)$ den Bedingungen 23, c) und 24, c). Man sieht nun leicht, daß $J(x)$ sowohl ein $\psi_3(x)$ wie ein $\varphi_3(x)$ ist; $f(x)$ ist P-integrierbar im Sinne der Definition 25, und beide Integrationen liefern denselben Wert.

§ 11.

Schließlich läßt sich ein P-Integral $S_3^*(x)$ einführen von gleicher Tragweite wie $S_3(x)$, jedoch mittels adjungierter Funktionen, bei deren Definition die Begriffe UTV und OTV nicht gebraucht werden, wodurch eine engere Anschließung an die Perron-Bauersche Integraldefinition erreicht wird.

Definition 26. Eine *zu* $f(x)$ *in* (a, b) *adjungierte Oberfunktion* $\psi_3^*(x)$ soll, außer stetig zu sein für $a \leqq x \leqq b$, den weiteren Bedingungen genügen: a) $\psi_3^*(a) = 0$; b) das Intervall (a, b) läßt sich in höchstens abzählbar unendlich viele Teilmengen (H_j) zerlegen, derartig, daß für jedes j auf dem Kerne K_j der ersten Ableitung H_j' (falls dieser Kern existiert) die

---

[49]) Siehe Denjoy, loc. cit. [4]), zweites Zitat, S. 184; erstes Zitat, S. 173 und 175 oder unsere in [35]) zitierte Arbeit, th. 12, 14, 9.

unteren Derivierten von $\psi_3^*(x)$ „in bezug auf dieses K_j" $\neq -\infty$ und fast überall $\geqq f(x)$ sind[50]).

Definition 27. Eine *zu* $f(x)$ *in* (a,b) *adjungierte Unterfunktion* $\varphi_3^*(x)$ ist stetig für $a \leqq x \leqq b$ und genügt weiter den Bedingungen: a) $\varphi_3^*(a) = 0$; b) das Intervall (a,b) läßt sich in höchstens abzählbar unendlich viele Teilmengen (H_j) zerlegen, derartig, daß für jedes j auf dem Kerne K_j der ersten Ableitung H_j' (falls dieser Kern existiert) die oberen Derivierten von $\varphi_3^*(x)$ „in bezug auf dieses K_j" $\neq +\infty$ und fast überall $\leqq f(x)$ sind[50]).

Satz XXIX. *Jedes* $\psi_3^*(x)$ *ist UTV, und jedes* $\varphi_3^*(x)$ *ist OTV in* (a,b).

Beweis. Betrachten wir ein $\psi_3^*(x)$. Die Summe derjenigen Mengen (H_j), deren erste Ableitung verschwindet oder höchstens abzählbar ist, wird auch abzählbar sein; jeder ihrer Punkte ist aufzufassen als eine Menge, auf der $\psi_3^*(x)$ UT ist.

Wenn für eine Menge H_j der Kern K_j von H_j' nicht verschwindet, so wird eine stetige Funktion $g_j(x)$, welche auf K_j und in a und b mit $\psi_3^*(x)$ zusammenfällt und sich in den Intervallen der in bezug auf (a,b) komplementären Menge $C(K_j)$ linear ändert, UTV* sein in (a,b); dies läßt sich zeigen nach dem Beweisverfahren von § 9. Daraus folgt, daß sich K_j in höchstens abzählbar unendlich viele Teilmengen zerlegen läßt, auf denen $\psi_3^*(x)$ UT sein wird.

Man sieht nun leicht, daß $\psi_3^*(x)$ in (a,b) UTV ist.

Definition 28. Wir definieren für die in (a,b) gegebene Funktion $f(x)$ *das unbestimmte obere P-Integral*

$$\Psi_3^*(x) \equiv \overline{\int_a^x} (P_3^*)\, f(x)\, dx = \operatorname{fin\,inf} \psi_3^*(x)$$

und *das unbestimmte untere P-Integral*

$$\Phi_3^*(x) \equiv \underline{\int_a^x} (P_3^*)\, f(x)\, dx = \operatorname{fin\,sup} \varphi_3^*(x).$$

Wenn in (a,b) $\Psi_3^*(x) = \Phi_3^*(x)$ ist, so definiert ihr gemeinsamer Wert *das unbestimmte P-Integral* $S_3^*(x)$ in (a,b).

Satz G. *Die Definition von* $S_3^*(x)$ *ist äquivalent mit der Definition des allgemeinen D-Integrals.*

Beweis. Aus den Definitionen 23, 24, 26, 27 und Satz XXIX geht hervor, daß die Klassen der $\{\psi_3^*(x)\}$ und $\{\varphi_3^*(x)\}$ in den Klassen der $\{\psi_3(x)\}$ und $\{\varphi_3(x)\}$ enthalten sind. Daraus folgt (siehe Def. 25 und 28), daß, wenn eine Funktion ein $S_3^*(x)$ hat, dieses Integral auch ein $S_3(x)$ ist und nach Satz F ebenfalls ein allgemeines D-Integral.

---

[50]) Wenn also H_j' abzählbar ist, wird keinerlei Bedingung gefordert.

Umgekehrt, es existiere ein allgemeines D-Integral $\int_a^b (D) f(x)\,dx$. Bei seiner Konstruktion bestimmt man eine wohlgeordnete Folge von perfekten Mengen, deren Anzahl abzählbar ist und welche u. a. folgende Eigenschaft haben. Wenn für eine dieser Mengen, P_α, die in bezug auf das Intervall $(a \leqq x \leqq b)$ komplementäre Menge $C(P_\alpha)$ gebildet wird von den offenen Intervallen $(a_j^{(\alpha)}, b_j^{(\alpha)})$, so wird es eine auf P_α nirgends dichte, abgeschlossene Teilmenge $M_{\alpha+1}$ geben, derartig, daß jeder Punkt ξ von $P_\alpha - M_{\alpha+1}$ innerer Punkt eines abgeschlossenen Intervalls (c, d) ist, dessen Endpunkte zu P_α gehören, für das $\int_{(c,d)\cdot P_\alpha} (L) f(x)\,dx$ existiert, während die Summe der über den in ihm liegenden Intervallen $(a_j^{(\alpha)}, b_j^{(\alpha)})$ genommenen D-Integralen $\sum_{(c,d)} \int_{a_j^{(\alpha)}}^{b_j^{(\alpha)}} (D) f(x)\,dx$ eine absolut konvergente Reihe ist; $P_{\alpha+1}$ ist der perfekte Kern von $M_{\alpha+1}$.

Wir nehmen an, daß es möglich sei, für jedes im Innern eines derartigen $(a_j^{(\alpha)}, b_j^{(\alpha)})$ liegende, abgeschlossene Intervall (p, q) zu $f(x)$ adjungierte Oberfunktionen $\psi_3^{*(p,q)}(x)$ zu bestimmen im Sinne der Definition 26, welche größer sind als, jedoch beliebig wenig abweichen von dem (in diesem Intervall existierenden) allgemeinen D-Integral. Da

$$\int_{a_j^{(\alpha)}}^{b_j^{(\alpha)}} (D) f(x)\,dx = \lim_{k=\infty} \int_{a_{j,k}^{(\alpha)}}^{b_{j,k}^{(\alpha)}} (D) f(x)\,dx$$

ist, wobei $a_j^{(\alpha)} < a_{j,k+1}^{(\alpha)} < a_{j,k}^{(\alpha)} < b_{j,k}^{(\alpha)} < b_{j,k+1}^{(\alpha)} < b_j^{(\alpha)}$ $(k = 0, 1, 2, \ldots)$; $a_{j,0}^{(\alpha)} = b_{j,0}^{(\alpha)} = \frac{1}{2}(a_j^{(\alpha)} + b_j^{(\alpha)})$; $\lim_{k=\infty} a_{j,k}^{(\alpha)} = a_j^{(\alpha)}$ und $\lim_{k=\infty} b_{j,k}^{(\alpha)} = b_j^{(\alpha)}$ sein wird, so muß es bei willkürlich positivem ε möglich sein, die Punkte $a_{j,1}^{(\alpha)}, a_{j,2}^{(\alpha)}, \ldots$ und $b_{j,1}^{(\alpha)}, b_{j,2}^{(\alpha)}, \ldots$ so zu wählen, daß die Schwankung des unbestimmten D-Integrals in jedem der Intervalle $(a_j^{(\alpha)}, a_{j,k}^{(\alpha)})$ und $(b_{j,k}^{(\alpha)}, b_j^{(\alpha)})$ $(k \geqq 1)$ kleiner als $\frac{\varepsilon}{2^k}$ sei.

In einem Intervall $(a_{j,k+1}^{(\alpha)}, a_{j,k}^{(\alpha)})$ $(k \geqq 0)$ existiert eine zu $f(x)$ adjungierte Oberfunktion $\psi_3^{*(a_{j,k+1}^{(\alpha)}, a_{j,k}^{(\alpha)})}(x)$, deren Werte größer sind als, jedoch um weniger als $\frac{\varepsilon}{2^{k+1}}$ abweichen von den übereinstimmenden Werten von $\int_{a_{j,k+1}^{(\alpha)}}^{x} (D) f(x)\,dx$; auch in einem Intervall $(b_{j,k}^{(\alpha)}, b_{j,k+1}^{(\alpha)})$ $(k \geqq 0)$ existiert ein $\psi_3^{*(b_{j,k}^{(\alpha)}, b_{j,k+1}^{(\alpha)})}(x)$, größer als, jedoch um weniger als $\frac{\varepsilon}{2^{k+1}}$ abweichend von $\int_{b_{j,k}^{(\alpha)}}^{x} (D) f(x)\,dx$.

Es läßt sich unschwer zeigen, daß

16) $$\sum_{k=0}^{\infty} \psi_3^{*(a_{j,k+1}^{(\alpha)}, a_{j,k}^{(\alpha)})}(a_{j,k}^{(\alpha)}) \quad \text{und} \quad \sum_{k=0}^{\infty} \psi_3^{*(b_{j,k}^{(\alpha)}, b_{j,k+1}^{(\alpha)})}(b_{j,k+1}^{(\alpha)})$$

konvergente Reihen sind.

Wir definieren nun eine Funktion, gleich Null in $a_j^{(\alpha)}$, gleich der Summe der beiden Reihen (16) in $b_j^{(\alpha)}$, für $a_{j,n+1}^{(\alpha)} \leqq x < a_{j,n}^{(\alpha)}$ $(n \geqq 0)$ gleich

$$\sum_{k=n+1}^{\infty} \psi_3^{*(a_{j,k+1}^{(\alpha)}, a_{j,k}^{(\alpha)})}(a_{j,k}^{(\alpha)}) + \psi_3^{*(a_{j,n+1}^{(\alpha)}, a_{j,n}^{(\alpha)})}(x)$$

und für $b_{j,n}^{(\alpha)} \leqq x < b_{j,n+1}^{(\alpha)}$ $(n \geqq 0)$ gleich

$$\sum_{k=0}^{\infty} \psi_3^{*(a_{j,k+1}^{(\alpha)}, a_{j,k}^{(\alpha)})}(a_{j,k}^{(\alpha)}) + \sum_{k=0}^{n-1} \psi_3^{*(b_{j,k}^{(\alpha)}, b_{j,k+1}^{(\alpha)})}(b_{j,k+1}^{(\alpha)}) + \psi_3^{*(b_{j,n}^{(\alpha)}, b_{j,n+1}^{(\alpha)})}(x).$$

Die auf diese Weise für $a_j^{(\alpha)} \leqq x \leqq b_j^{(\alpha)}$ definierte Funktion ist eine in $(a_j^{(\alpha)}, b_j^{(\alpha)})$ zu $f(x)$ adjungierte Oberfunktion $\psi_3^{*(a_j^{(\alpha)}, b_j^{(\alpha)})}(x)$ (im Sinne der Definition 26), welche in jedem Punkt von $(a_j^{(\alpha)}, b_j^{(\alpha)})$ größer ist als $\int_{a_j^{(\alpha)}}^{x} (D) f(x)\, dx$, aber um weniger als 2ε davon abweicht.

Eine für $c \leqq x \leqq d$ definierte Funktion $g(x)$, gleich $f(x)$ auf $(c, d) \cdot P_\alpha$ und gleich Null in den komplementären Intervallen $(a_j^{(\alpha)}, b_j^{(\alpha)})$, ist L-integrierbar. Nach de la Vallée Poussin[51]) existiert für $c \leqq x \leqq d$ eine Majorante $\mathfrak{M}^{(\alpha; c, d)}(x)$, welche für jedes x $\int_c^x (L) g(x)\, dx$ um weniger als η übertrifft (η willkürlich positiv) und deren untere Derivierte für jedes x[52]) $\neq -\infty$ und $\geqq g(x)$ sind.

In jedem Intervall $(a_j^{(\alpha)}, b_j^{(\alpha)})$ sind nach dem Vorherigen die Oberfunktionen $\psi_3^{*(a_j^{(\alpha)}, b_j^{(\alpha)})}(x)$ so zu wählen, daß sie von $\int_{a_j^{(\alpha)}}^{x} (D) f(x)\, dx$ um weniger als $\frac{\eta}{2^j}$ abweichen. Es läßt sich dann eine ganze Zahl J angeben, derartig, daß

$$\sum_{j=J}^{\infty} \left| \psi_3^{*(a_j^{(\alpha)}, b_j^{(\alpha)})}(b_j^{(\alpha)}) \right| < \eta$$

klein wird.

In den abgeschlossenen Intervallen $(a_j^{(\alpha)}, b_j^{(\alpha)})$ definieren wir nun Funktionen $g_j^{(\alpha)}(x)$; wir nehmen für jedes $j < J$

$$g_j^{(\alpha)}(x) = \psi_3^{*(a_j^{(\alpha)}, b_j^{(\alpha)})}(x),$$

---

[51]) Siehe z. B. de la Vallée Poussin, Intégrales de Lebesgue, etc., S. 74 – 76.

[52]) Hieraus folgt, daß die im folgenden konstruierten Oberfunktionen den Bedingungen der Definition 26 auch genügen, wenn man in dieser Definition „fast überall" aus der Bedingung b) fortläßt. Damit ist dann die in der Einleitung gegebene Modifikation des Satzes G gerechtfertigt.

für jedes der übrigen Intervalle

$$g_j^{(\alpha)}(x) = \psi_3^{*(a_j^{(\alpha)}, b_j^{(\alpha)})}(x) + \frac{x - a_j^{(\alpha)}}{b_j^{(\alpha)} - a_j^{(\alpha)}} \cdot \left| \psi_3^{*(a_j^{(\alpha)}, b_j^{(\alpha)})}(b_i^{(\alpha)}) \right|.$$

Definieren wir schließlich für $c \leqq x \leqq d$

$$\psi_3^{*(c,d)}(x) = \mathfrak{M}^{(\alpha; c, d)}(x) + \sum_{(j)}^{x} g_j^{(\alpha)}(b_j^{(\alpha)}) + g_I^{(\alpha)}(x),$$

wobei die Summation erstreckt wird über diejenigen abgeschlossenen Intervalle $(a_j^{(\alpha)}, b_j^{(\alpha)})$, für die $b_j^{(\alpha)} < x$ ist, während $g_I^{(\alpha)}(x)$ gleich Null zu rechnen ist, wenn x nicht zu einem abgeschlossenen Intervall $(a_I^{(\alpha)}, b_I^{(\alpha)})$ gehört. Diese Funktion ist nun in (c, d) eine zu $f(x)$ adjungierte Oberfunktion im Sinne der Definition 26, und es ist in (c, d)

$$0 \leqq \psi_3^{*(c,d)}(x) - \int_c^x (D) f(x)\, dx < 3\eta.$$

Man sieht nun leicht, daß es in abgeschlossenen Intervallen, deren innere Punkte zu der in bezug auf (a, b) komplementären Menge $C(M_{\alpha+1})$ gehören, Oberfunktionen ψ_3^* gibt, welche beliebig wenig vom unbestimmten allgemeinen D-Integral abweichen. Dasselbe gilt also auch für abgeschlossene Intervalle mit inneren Punkten, gehörend zu der in bezug auf (a, b) komplementären Menge des perfekten Kerns $P_{\alpha+1}$ von $M_{\alpha+1}$.

Es sei M_1 die Menge der Punkte von (a, b), in deren jede Umgebung $f(x)$ nicht L-integrierbar ist. Dann lassen sich, ausgehend von den Majoranten von de la Vallée Poussin, in den abgeschlossenen, zu M_1 komplementären Intervallen Oberfunktionen ψ_3^* konstruieren, welche beliebig wenig größer sind als das unbestimmte D-Integral. Derartige Oberfunktionen existieren somit auch für die abgeschlossenen Intervalle, welche komplementär sind zu dem perfekten Kerne P_1 von M_1. Da es eine Ordnungszahl der ersten oder zweiten Zahlenklasse gibt, für die P_α leer ist, so führt das Vorhergehende, mittels transfiniter Induktion, zur Existenz von Oberfunktionen $\psi_3^*(x)$ für (a, b), welche beliebig wenig von $\int_a^x (D) f(x)\, dx$ abweichen[52a]).

[52a]) Nachträgliche Bemerkung. Wenn man ausgeht von der deskriptiven Definition des allgemeinen D-Integrals $J(x)$ von $f(x)$ (siehe den Anfang des Beweises von Satz F), so läßt sich eine Oberfunktion $\psi_3^*(x)$, welche in (a, b) um weniger als ε (ε beliebig positiv) von $J(x)$ abweicht, auch auf folgende Weise angeben. Aus dieser Definition geht mittels transfiniter Induktion hervor, daß sich (a, b), ausgenommen in den Punkten einer höchstens abzählbaren Menge E, überdecken läßt von endlich oder abzählbar unendlich vielen perfekten Mengen (P_k), auf deren jeder $J(x)$ totalstetig ist. Zu jeder Menge P_k wähle man ein positives ε_k, derartig, daß $\varepsilon_1 + \varepsilon_2 + \ldots \varepsilon_k + \ldots = \varepsilon$ sei. Jede stetige Funktion $J_k(x)$, welche in a und b und in den Punkten des zugehörigen P_k dieselben Werte hat wie das unbestimmte D-Integral $J(x)$ und welche sich in den zu P_k komplementären Intervallen von (a, b) linear

(Fortsetzung der Fußnote 52a auf nächster Seite.)

Obige Konstruktion, angewandt für die Funktion $-f(x)$, führt zu Oberfunktionen, deren entgegengesetzte Werte in (a, b) zu $f(x)$ adjungierte Unterfunktionen liefern im Sinne der Definition 27, welche beliebig wenig von $\int_a^x (D) f(x)\,dx$ abweichen. Daher hat $f(x)$ auch ein unbestimmtes P-Integral $S_3^*(x)$ in (a, b), gleich $\int_a^x (D) f(x)\,dx$.[53])

§ 12.

Die Definitionen 25 und 28 und die in der Einleitung gegebene Integraldefinition führen ebensoweit. Nennen wir ein Integral, gebildet nach einer dieser Definitionen, *ein allgemeines P-Integral*. Für dieses allgemeine P-Integral lassen sich wieder viele der bekannten Integralsätze beweisen, wobei der Beweis auf dieselbe einfache Weise verläuft wie beim Perron-Bauerschen Integral[54]). Nur zwei Sätze wollen wir völlig beweisen:

ändert, ist totalstetig in (a, b); es existiert somit eine in (a, b) nicht-abnehmende, stetige Funktion $\chi_k(x)$ mit $\chi_k(a) = 0$ und $\chi_k(b) < \varepsilon_k$, welche, zu $J_k(x)$ addiert, eine Majorante von $J_k(x) = \int_a^x (L) J_k'(x)\,dx$ liefert im de la Vallée-Poussinschen Sinne. Man sieht nun leicht, daß $J(x) + \sum_{(k)} \chi_k(x)$ eine Oberfunktion $\psi_3^*(x)$ von $J(x)$ sein wird, welche in (a, b) um weniger als ε von $J(x)$ abweicht.

[53]) *Die Definitionen* 25 *und* 28 *führen zu Integralbegriffen äquivalent mit dem speziellen Denjoyschen Integralbegriff, wenn man in diese Definitionen eine Bedingung hinzufügt*: „*Es existiere mindestens eine Oberfunktion* $\psi_3(x)$ *bzw.* $\psi_3^*(x)$, *welche OTV* sei in* (a, b), *und mindestens eine Unterfunktion* $\varphi_3(x)$ *bzw.* $\varphi_3^*(x)$, *welche UTV* sei in* (a, b)." Vgl. auch Fußn. [24]).

Sie führen zu Integralbegriffen äquivalent mit dem Denjoy-Khintchineschen, wenn man in beide die Bedingung hinzufügt: „*Es existiere mindestens eine Oberfunktion* $\psi_3(x)$ *bzw.* $\psi_3^*(x)$, *deren (allgemeine) obere Derivierte fast überall* $\neq +\infty$, *und mindestens eine Unterfunktion* $\varphi_3(x)$ *bzw.* $\varphi_3^*(x)$, *deren (allgemeine) untere Derivierte fast überall* $\neq -\infty$ *seien.*" Vgl. hiermit auch Satz B (§ 3).

Ein Teil des Beweises der letzten Behauptung fordert Anwendung des von Denjoy herrührenden Satzes (siehe z. B. Hobson, loc. cit. [2]), § 468):

In jedem komplementären Intervall (a_n, b_n) einer perfekten Menge P von positivem Maße sei eine Funktion $y_n(x)$ definiert, welche in a_n gleich Null und in b_n gleich V_n sei, während die oberen Schranken ω_n der $|y_n(x)|$ eine konvergente Reihe bilden. Wenn a und b die Endpunkte von P sind, definieren wir in (a, b) eine Funktion

$$f(x) = \sum_{(n)}^{x} V_n + y_m(x),$$

wobei die Glieder der Summe sich beziehen auf diejenigen Intervalle (a_n, b_n), für die $b_n < x$ ist, und $y_m(x)$ gleich Null zu rechnen ist, wenn x nicht zu einem abgeschlossenen Intervall (a_m, b_m) gehört. $f(x)$ hat dann fast überall eine Nullableitung auf P.

[54]) Siehe die von Kamke, loc. cit. [1]), § 14 gegebenen Sätze und Beweise, die sich alle unmittelbar beim allgemeinen P-Integral übertragen lassen.

a) *Eine in einem Intervall (a, b) ihr Vorzeichen nicht wechselnde Funktion $f(x)$ hat daselbst gleichzeitig ein L-Integral und ein allgemeines P-Integral von gleichem Werte.*

Beweis. Nehmen wir $f(x) \geqq 0$ an. Da $\int_a^x (L) f(x)\,dx$ sowohl ein ψ_3 wie ein φ_3 ist (siehe Def. 23 und 24), so ist dieses Integral, falls es existiert, zugleich das allgemeine P-Integral.

Umgekehrt, es existiere ein allgemeines P-Integral $\int_a^x (P) f(x)\,dx$, und $\psi(x)$ sei eine in (a, b) zu $f(x)$ adjungierte Oberfunktion gemäß der Definition der Einleitung. Dann folgt aus dem Hilfssatze von § 6, daß es eine in (a, b) nirgends dichte, abgeschlossene Menge M_1 gibt, in deren komplementären Intervallen die unteren Derivierten von $\psi(x) \geqq f(x)$ sind. Daraus folgt, daß $\psi(x)$ nicht-abnehmend ist in diesen Intervallen und somit auch in den komplementären Intervallen des perfekten Kerns P_1 von M_1. Aus dem Hilfssatze folgt weiter, daß P_1 eine auf P_1 nirgends dichte, abgeschlossene Menge M_2 enthält, derartig, daß jeder Punkt von $P_1 - M_2$ innerer Punkt eines Intervalls (c, d) ist mit der Eigenschaft, daß auf $(c, d) \cdot P_1$ die in bezug auf diese Menge gebildeten unteren Derivierten von $\psi(x) \geqq f(x)$ sind. Da $\psi(x)$ in den komplementären Intervallen von $(c, d) \cdot P_1$ nicht-abnehmend ist, so sind in (c, d) die allgemeinen unteren Derivierten von $\psi(x) \geqq 0$. Also wird $\psi(x)$ auch nicht-abnehmend sein in (c, d), schließlich in jedem komplementären Intervall des perfekten Kerns P_2 von M_2. Das Verfahren transfinit fortsetzend, folgt, daß auch in (a, b) $\psi(x)$ nicht-abnehmend ist. Da die Oberfunktion $\psi(x)$ übrigens willkürlich ist, so gibt Anwendung des Satzes XXVII dieselbe Eigenschaft für das allgemeine P-Integral in (a, b). Dieses Integral besitzt somit fast überall in (a, b) eine endliche, summierbare Ableitung, welche nach Satz XXVIII fast überall gleich $f(x)$ ist. Also muß

$$\int_a^x (P) f(x)\,dx = \int_a^x (L) f(x)\,dx$$

sein [55]).

---

[55]) Vgl. Denjoy, loc. cit. [4]), zweites Zitat, S. 202. Es läßt sich aus obigem Satze folgern (vgl. Denjoy, loc. cit.): Wenn die Funktionen $f_n(x)$ in (a, b) ein allgemeines P-Integral haben und bei zunehmendem n wachsend gegen eine Funktion $f(x)$ konvergieren, so folgt aus der Endlichkeit von $\lim_{n=\infty} \int_a^b (P) f_n(x)\,dx$ die Existenz des allgemeinen P-Integrals $\int_a^b (P) f(x)\,dx$, und umgekehrt. Überdies wird dann

$$\lim_{n=\infty} \int_a^b (P) f_n(x)\,dx = \int_a^b (P) f(x)\,dx$$

sein.

b) *Es existiere das allgemeine P-Integral über* (a, b) *für jede der Funktionen* $f_1(x), f_2(x), \ldots, f_n(x), \ldots$, *und die Funktionen mögen in* (a, b) *eine Grenzfunktion* $f(x)$ *haben. Weiter sei in* (a, b) *für jedes* n

$$|f(x) - f_n(x)| \leqq G(x),$$

wobei $G(x)$ *eine über* (a, b) *summierbare, nicht-negative Funktion darstellt.*

Dann wird man aus der Existenz von $\int_a^b (P) f(x)\,dx$ *folgern können:*

$$\int_a^b (P) f(x)\,dx = \lim_{n=\infty} \int_a^b (P) f_n(x)\,dx.$$

Beweis. Die Funktion $f(x) - f_n(x)$ ist meßbar und ihr absoluter Wert $\leqq G(x)$. Daraus folgt die Summierbarkeit von $f(x) - f_n(x)$, und nach einem Lebesgueschen Satze wird

$$\lim_{n=\infty} \int_a^b (L)\{f(x) - f_n(x)\}\,dx = \int_a^b (L) \lim_{n=\infty} \{f(x) - f_n(x)\}\,dx = 0$$

sein. Daher ist auch

$$\lim_{n=\infty} \left\{ \int_a^b (P) f(x)\,dx - \int_a^b (P) f_n(x)\,dx \right\} = \lim_{n=\infty} \int_a^b (P)\{f(x) - f_n(x)\}\,dx = 0.$$

27. September 1930.

(Eingegangen am 1. Oktober 1930.)

The Approximately Continuous Perron Integral.

Von

J. C. Burkill in Cambridge.

§ 1.

Introduction.

From a given function $f(x)$ satisfying certain conditions, Denjoy's first (special) definition of the integral constructs a continuous function $F(x)$ of which $f(x)$ is almost everywhere the derivative. His later (general) process constructs a continuous $F(x)$, starting from a function $f(x)$ satisfying less restrictive conditions. In order to extend the fundamental theorem of differentiation to the general integral, Denjoy and Khintchine showed that it is necessary to generalize the idea of derivative to that of approximate derivative. It is then true that $F(x)$ has almost everywhere an approximate derivative equal to $f(x)$. If it is given that $f(x)$ is at all points the finite approximate derivative of a continuous function, then Denjoy's process will yield this continuous function, which is determinate except for an additive constant.

We observe, however, that a function which has everywhere a finite approximate derivative need not be continuous; all we can say is that it must be approximately continuous. When an integral has been evolved of such generality that it is expedient to consider its approximate derivative rather than its ordinary derivative, we may ask whether Denjoy's restriction of the integral to be continuous is quite natural and whether we should not instead postulate only approximate continuity. In this paper we define an approximately continuous integral and establish its most straightforward properties. We adopt Perron's method of definition[1]) which seems to offer the simplest line of approach.

---

[1]) O. Perron, Sitzungsberichte der Heidelberger Akademie der Wissenschaften, Math.-naturw. Klasse, Abt. A (1914), 14. Abhandlung. An account of the Perron integral with further references is given by E. Kamke, Das Lebesguesche Integral, Teubner 1925.

Integrals which are not necessarily continuous have been investigated by Young[2]) — he calls them Y-integrals. He does not single out approximately continuous functions for special consideration and his definition is so wide that the integral does not possess the usual elementary properties and so is not of service in analysis. With Young's definition, the sum of two integrable functions is not necessarily integrable. Further, I am not aware that his integral has any differential properties.

We give now a simple illustration of the effect of postulating approximate continuity instead of continuity in the integral.

For an interval (a, b) define a continuously differentiable function $\varphi(k, a, b; x)$, zero outside the interval, and within the interval giving a dome-shaped curve of height k; we may take, say, $\varphi = k \sin^2\{(x-a)/(b-a)\}$.

On the x-axis take a sequence of points

$$a_1 < b_1 \leqq a_2 < \ldots \leqq a_n < b_n \ldots \to c,$$

such that the set of points belonging to the intervals (a_n, b_n) has density zero at the limit point c.

Let

$$\Phi(x) = \sum_{n=1}^{\infty} \varphi(k_n, a_n, b_n; x)$$

where k_n does not tend to zero (the sequence k_n may be unbounded).

Then $\Phi(x)$ is continuous and has a finite derivative for every value of x except c, where it is approximately continuous and has approximate derivative zero. The derivative (or approximate derivative) is integrable by Denjoy's process in any interval not containing c, but, on account of the discontinuity of $\Phi(x)$ at c, is not integrable in any interval containing c. Our extension of Perron's method will integrate the derivative in any interval, giving the primitive function $\Phi(x)$.

§ 2.

The upper and lower approximate derivates.

We define the (two-sided) upper approximate derivate of a measurable function $f(x)$ for a given value of x to be the lower bound of numbers a such that the set of values of ξ for which

$$\frac{f(\xi)-f(x)}{\xi-x} \geqq a$$

has zero density at x. We write this derivate as

$$AD^*f(x) \quad \text{or} \quad AD^*f.$$

---

[2]) W. H. Young, Proc. London Math. Soc. (2) 16 (1917), pp. 175–218.

The lower approximate derivate AD_*f is defined in a corresponding way. When $AD^*f = AD_*f$, f has an approximate derivative ADf.

We may note the relation between these derivates and the one-sided derivates of measurable functions with which Dr. Haslam-Jones and I have previously worked[3]). We defined the upper and lower right and the upper and lower left approximate derivates, written respectively as AD^+f, AD_+f, AD^-f, AD_-f; and we showed that they are measurable functions. We have now

$$AD^*f = \max(AD^+f, AD^-f),$$
$$AD_*f = \min(AD_+f, AD_-f)$$

and it follows that AD^*f and AD_*f are measurable.

The following properties are immediate

$$AD_*f \leqq AD^*f,$$
$$AD_*(-f) = -AD^*f,$$
$$AD^*(f_1+f_2) \leqq AD^*f_1 + AD^*f_2$$
$$AD_*(f_1+f_2) \geqq AD_*f_1 + AD_*f_2.$$

We need also a lemma.

Lemma. *If, for* $a \leqq x \leqq b$,

(1) $f(x)$ *is approximately continuous,*

(2) $AD_*f(x) \geqq 0$ [4]),

then

$$f(b) \geqq f(a).$$

Fix k where $0 < k < 1$ and take ε (small).

Since $AD_*f(a) \geqq 0$, the set of x for which

(A) $$f(x) - f(a) \geqq -\varepsilon(x-a)$$

has unit density at a. We can therefore take x_1, arbitrarily near to a, such that the inequality (A) is true for $x = x_1$ and also for a set of x in (a, x_1) of measure at least $k(x_1 - a)$.

Starting from x_1, define x_2 in the same way that x_1 was obtained from a. We thus generate a Lebesgue chain of intervals. Suppose that the x_n tend to a limit-point ξ. The inequality (A) holds for a set of x having average density at least k in each interval (x_{n-1}, x_n). Since $f(x)$ is approximately continuous at ξ, it follows that

$$f(\xi) - f(a) \geqq -\varepsilon(\xi - a).$$

---

[3]) J. C. Burkill and U. S. Haslam-Jones, Proc. London Math. Soc. (2) **32** (1931).

[4]) The same argument will hold if condition (2) is replaced by $AD_+f(x) \geqq 0$ or by $AD_-f(x) \geqq 0$.

Covering the interval (a, b) with the Lebesgue chain we have

$$f(b) - f(a) \geqq -\varepsilon(b-a)$$

nd, since ε is arbitrary,

$$f(b) \geqq f(a).$$

It is easy to see, further, that an approximately continuous function hich is monotonic must be continuous, so that $f(x)$ is continuous for $< x < b$.

§ 3.

Major and minor functions and definition of the integral.

Suppose that $f(x)$ is measurable and finite almost everywhere in $a, b)$. We say that $M(x)$ is a major function of $f(x)$ in (a, b) if

(1) $M(x)$ is approximately continuous in (a, b),

(2) $M(a) = 0$,

nd, for all x in $a \leqq x \leqq b$,

(3) $AD_*M(x) > -\infty$,

(4) $AD_*M(x) \geqq f(x)$.

A minor function $m(x)$ is defined in a corresponding way.

Lemma. *If $M(x)$ is any major function and $m(x)$ any minor unction of $f(x)$, then $M(x) - m(x)$ is an increasing function and in articular $M(b) \geqq m(b)$.*

For

$$\begin{aligned} AD_*\{M(x) - m(x)\} &\geqq AD_*M(x) + AD_*\{-m(x)\} \\ &= AD_*M(x) - AD^*m(x) \geqq 0 \end{aligned}$$

emembering that $AD_*M(x) > -\infty$ and $AD^*m(x) < +\infty$)

The result now follows from the Lemma of § 2.

If $f(x)$ has major and minor functions, then

$$I(b) = \text{lower bound of all } M(b),$$

$$J(b) = \text{upper bound of all } m(b)$$

e finite and $I(b) \geqq J(b)$.

If $I(b) = J(b)$ we say that $f(x)$ has an *approximately continuous erron integral*

$$(AP)\int_a^b f(x)\,dx$$

ual to their common value. More generally we may call $I(b)$ and $J(b)$ e upper and lower AP-integrals. In future we shall usually omit the

distinguishing mark (AP) in front of the integral and in the absence of a special remark we shall understand integration to be taken in this sense.

We now investigate the effect of modifying the definitions of major and minor functions. Let us say that $M_1(x)$ is a generalized major function if it satisfies the conditions (1), (2), (3) laid down for $M(x)$ and, instead of (4),

$(4')$ $\quad AD_* M_1(x) \geqq f(x)$ almost everywhere in (a, b).

We shall prove that replacing (4) by $(4')$ leaves the scope of the integral unaltered.

Since every $M(x)$ is an $M_1(x)$, we have

$$\text{lower bound of all } M_1(b) \leqq \text{lower bound of all } M(b).$$

We have to prove that these bounds are equal.

Lemma. *Given a set E of measure zero in (a, b) and any positive ε, we can construct a continuous increasing function $\mu(x)$ such that*

$$\mu(a) = 0, \quad \mu(b) < \varepsilon,$$
$$\mu'(x) = \infty \text{ for } x \text{ in } E.$$

For, given n, there is an open set O_n, containing E, of measure less than $\varepsilon/4^n$.

Write

$$\nu_n(x) = 2^n \text{ for } x_n \text{ in } O_n,$$
$$= 0 \text{ for other } x,$$

and let

$$\mu_n(x) = \int_a^x \nu_n(x)\, dx.$$

Then $\mu_n(x)$ is continuous, increasing and $\mu_n(b) < \varepsilon/2^n$.

It is easy to see that $\mu(x) = \sum_{n=1}^{\infty} \mu_n(x)$ satisfies the conditions required in the Lemma.

We now prove that, if $f(x)$ has a generalized major function $M_1(x)$, then there is a major function $M(x)$ according to the original definition such that

$$M(b) < M_1(b) + \varepsilon.$$

For let E be the set of measure zero in which the inequality

$$AD_* M_1(x) \geqq f(x)$$

does not hold.

With the function $\mu(x)$ of the lemma, take

$$M(x) = M_1(x) + \mu(x).$$

Then the first three conditions for $M(x)$ are satisfied, and, since

$$AD_* M(x) \geqq AD_* M_1(x) + \mu'(x),$$

so is the fourth and our result is proved.

We may therefore adopt (1), (2), (3), (4′) as the standard conditions for a major function and we may drop the suffix from $M_1(x)$ and the word *generalized*.

The proof we have given of the possibility of generalizing the original major functions establishes the following extension of the lemma of § 2.

If, for $a \leqq x \leqq b$,

(1) $f(x)$ *is approximately continuous,*

(2) $AD_* f(x) \geqq 0$ *almost everywhere,*

(3) $AD_* f(x) > -\infty$ *everywhere,*

then $f(b) \geqq f(a)$ *(and* $f(x)$ *is continuous for* $a < x < b$*).*

§ 4.

Properties of the Integral.

We state first some properties which follow very simply from the definition.

If $f(x)$ *is integrable in* (a, b)*, it is integrable in any subinterval.*

If $b < a$*, we define* $\int_a^b f(x)\,dx$ *to be* $-\int_b^a f(x)\,dx$.

$$\int_a^b f(x)\,dx + \int_b^c f(x)\,dx = \int_a^c f(x)\,dx,$$

if the integral exists in the greatest of the three intervals.

$$\int_a^b c f(x)\,dx = c \int_a^b f(x)\,dx.$$

$$\int_a^b \{f_1(x) + f_2(x)\}\,dx = \int_a^b f_1(x)\,dx + \int_a^b f_2(x)\,dx,$$

if the right-hand side exists.

If $f_1(x) \leqq f_2(x)$*, then* $\int_a^b f_1(x)\,dx \leqq \int_a^b f_2(x)\,dx$.

If $f_1(x) = f_2(x)$ *almost everywhere and* $\int_a^b f_1(x)\,dx$ *exists, so does* $\int_a^b f_2(x)\,dx$ *and the two integrals are equal.*

If $f(x)$ is integrable in the ordinary Perron sense, (or what is equivalent, in the special Denjoy sense) then it is integrable according to our definition and the integrals have the same value.

We shall now establish some further properties.

Lemma. *If in (a, b) each $s_n(x)$ is approximately continuous and $s_n(x)$ tends uniformly to $s(x)$, then $s(x)$ is approximately continuous.*

Let c be a point of (a, b).

Given ε, choose n so that

$$|s(x) - s_n(x)| < \tfrac{1}{3}\varepsilon \quad \text{for all } x \text{ in } (a, b).$$

The set of points x for which

$$|s_n(x) - s_n(c)| > \tfrac{1}{3}\varepsilon$$

has zero density at c.

Therefore the set for which

$$|s(x) - s(c)| > \varepsilon$$

has zero density at c, and the lemma is proved.

I. $F(x) = \int_a^x f(x)\,dx$ *is an approximately continuous function of* x.

Given $\varepsilon_n \to 0$, there are major and minor functions such that

$$m_n(x) \leqq F(x) \leqq M_n(x),$$
$$0 \leqq M_n(x) - m_n(x) < \varepsilon_n.$$

Hence $F(x)$ is the limit of a uniformly convergent sequence of approximately continuous functions and the result follows from the Lemma.

II. *If, at all points of (a, b), $G(x)$ has a finite approximate derivative $g(x)$, then $g(x)$ is integrable and*

$$\int_a^b g(x)\,dx = G(b) - G(a).$$

For $G(x) - G(a)$ satisfies the conditions for a major function and those for a minor function.

III. *If $F(x) = \int_a^x f(x)\,dx$, then $ADF(x) = f(x)$ almost everywhere.*

If not, there is a positive number k and a set E of positive measure μ in which

$$\text{either } f - AD_*F > k \text{ or } AD^*F - f > k.$$

Suppose the former.

Take $\varepsilon < \frac{1}{2}\mu k$ and let $M(x)$ be a major function such that $M(x) - F(x)$, say $R(x)$, is less than ε for $a \leqq x \leqq b$.

Now $R(x)$ is an increasing function since $F(x)$ is the limit of minor functions $m_n(x)$ and each $M(x) - m_n(x)$ is an increasing function.

We have then

$$\int_a^b R'(x)\,dx \leqq R(b) - R(a) < \varepsilon,$$

the derivative existing almost everywhere.

From this inequality the set in which $R'(x) > \frac{1}{2}k$ has measure less than μ.

Therefore there is a set E_1 of positive measure contained in E in which $R'(x)$ exists and lies between 0 and $\frac{1}{2}k$.

At a point ξ of E_1,

$$f(\xi) \leqq AD_* M(\xi) = AD_* F(\xi) + R'(\xi) \leqq AD_* F(\xi) + \tfrac{1}{2}k.$$

which contradicts our original hypothesis.

IV. *If in* (a, b) $f(x) \geqq 0$ *and is integrable* (AP) *then* $f(x)$ *is integrable* (*Lebesgue*) *and the integrals in the two senses are equal.*

Let

$$F(x) = (AP)\int_a^x f(x)\,dx.$$

First, it is easy to see that $F(x)$ is an increasing (non-decreasing) function in (a, b). For if $M(x)$ is a major function for any subinterval (c, d) of (a, b), the first lemma of § 2 shows that $M(d) \geqq M(c)$. And, since F can be approximated to with arbitrary closeness by M's, it follows that $F(d) \geqq F(c)$.

Hence $F(x)$ has almost everywhere a derivative $F'(x)$ which is integrable (L), and

$$F(b) \geqq (L)\int_a^b F'(x)\,dx.$$

But Khintchine[5]) has shown that if an increasing function has at a given point a finite approximate derivative, it has an ordinary derivative of the same value.

Thus we have, combining the last inequality with the result of III, that $f(x)$ is integrable (L) and

$$F(b) \geqq (L)\int_a^b f(x)\,dx.$$

We have to prove that the equality sign must be taken. De la Vallée Poussin has shown that major functions (of a special type) may be constructed approximating with arbitrary closeness to the Lebesgue

[5]) A. Khintchine, Fundamenta Math. 9 (1927), pp. 212—279 (242).

integral. A major function in his sense serves as a major function for the AP-integral. Hence the inequality sign is impossible.

V. (*Limit under integral sign.*) *If*

(1) *for each* n, $g(x) \leqq f_n(x) \leqq h(x)$ *in* (a, b) *where* $g(x)$, $f_n(x)$, $h(x)$ *are integrable,*

(2) $f_n(x) \to f(x)$,

then $f(x)$ *is integrable, and*

$$\int_a^b f_n(x)\,dx \to \int_a^b f(x)\,dx.$$

Write

$$\varphi_n = f_n - g, \quad \varphi = f - g, \quad \psi = h - g.$$

Then φ_n and ψ are positive and integrable (AP).

By IV, they are integrable (L).

Since $0 \leqq \varphi_n \leqq \psi$, the Lebesgue theory of limits under the integral sign shows that φ is integrable (L) and

$$(L)\int_a^b \varphi_n\,dx \to (L)\int_a^b \varphi\,dx.$$

That is, the integrals now being (AP),

$$\int_a^b f_n\,dx - \int_a^b g\,dx \to \int_a^b f\,dx - \int_a^b g\,dx,$$

giving the result.

VI. (*Integration by parts.*) *Let* $F(x)$, $G(x)$ *have finite approximate derivatives* $f(x)$, $g(x)$ *in* (a, b). *If* $f(x)\,G(x)$ *is integrable, then* $F(x)\,g(x)$ *is integrable and*

$$\int_a^b Fg\,dx + \int_a^b fG\,dx = [FG]_a^b.$$

We have

$$f(x)\,G(x) + F(x)\,g(x) = AD\{F(x)\,G(x)\}$$

and the left-hand side is everywhere finite.

Using II, we have the result.

(Eingegangen am 5. November 1930.)

Bemerkungen zu der Zuschrift der Herren J. Bonder und S. Neumark in Band 32 dieser Zeitschrift[1]).

Von

W. Barth in Völklingen (Saar).

In der obigen Zuschrift werden die Bewegungen von Wirbeln in einer ruhenden Flüssigkeit bei verschiedenen eingebrachten Hindernissen untersucht. Dasselbe Problem wurde bereits in etwas anderer Form in einer Arbeit behandelt, die in der Zeitschrift für angewandte Mathematik und Mechanik veröffentlicht wurde[2]), und in welcher z. T. die gleichen Fälle durchgerechnet worden sind. In dieser Arbeit wurde für den ersten Fall das gleiche Ergebnis wie in obiger Zuschrift erhalten, dagegen zeigen die Lösungen für den zweiten Fall, der Wirbelströmung um ein geradliniges Hindernis von endlicher Länge, keine Übereinstimmung. Eine Nachfrage bei den Verfassern brachte die Klärung dieses Widerspruches dadurch, daß die Herren J. Bonder und S. Neumark, da sie als Ausgangsgebiet die Halbebene wählten, eine Zirkulation um das Hindernis angeordnet hatten, während in der anderen Arbeit diese absichtlich weggelassen wurde. Physikalisch betrachtet stellt daher die eine Lösung die Wirbelbewegung eines freien Wirbels um eine Platte mit angeordneter Zirkulationsströmung um die Platte, die andere Lösung die Wirbelbewegung ohne Zirkulation um die Platte dar. Die Verschiedenheit in den Grundannahmen bedingt auch einen ganz verschiedenartigen Verlauf der Wirbelbahnen insofern, als diese bei der letzten Annahme nach der Plattenmitte zu eingebuchtet verlaufen.

---

[1]) J. Bonder und S. Neumark, Einige Bemerkungen zu der Abhandlung „Über die Bewegung einzelner Wirbel in einer strömenden Flüssigkeit“ von M. Lagally, Math. Zeitschr. **32** (1930), S. 600—604.

[2]) W. Barth, Wirbelbahnen um Wände und Platten von unendlich kleiner Wandstärke, Zeitschrift für angew. Math. und Mech. **10** (1930), S. 247—251.

Vom physikalischen Standpunkte aus gibt es keinen entscheidenden Anlaß, einer der beiden Annahmen den Vorzug zu geben, da jeder der beiden Fälle sich in einiger Annäherung verwirklichen läßt. Für die Herren J. Bonder und S. Neumark ergab sich naturgemäß die erste Lösung, da für die Halbebene unter den gegebenen Bedingungen nur eine einzige Lösung existiert. — Diese ist vor allem dadurch ausgezeichnet, daß die Geschwindigkeit im Unendlichen mit $1/r^2$ abnimmt und damit die kinetische Energie außerhalb des Wirbelpunktes endlich bleibt. Dies trifft für die zweite Lösung nicht zu, jedoch ist diese ebenso berechtigt und dürfte vor allem für solche Fälle Anwendung finden, in denen ein Nullwerden der Zirkulation um das Hindernis gefordert werden muß.

(Eingegangen am 5. Januar 1931.)

Über den Permanenzsatz gewisser Limitierungsverfahren für Doppelfolgen.

Von

Friedrich Lösch in Stuttgart.

Die Übertragungen der regulären Limitierungsverfahren vom Toeplitzschen Typus (z. B. der C-Verfahren) auf Doppelfolgen sind bekanntlich im allgemeinen nicht mehr regulär. Der Grund dafür liegt einzig darin, daß bei Doppelfolgen die Konvergenz nicht wie bei einfachen Folgen die Beschränktheit nach sich zieht. Die genannten Verfahren erfüllen daher jedenfalls die Permanenzbedingung, wenn man nach dem Vorgang von Bromwich und Hardy nur beschränkte Folgen in Betracht zieht[1]).

Das Ziel der folgenden Zeilen ist nun, zu zeigen, daß gewisse, dem Toeplitzschen Typus entsprechende Limitierungsverfahren die Permanenzbedingung schon für eine größere Klasse von Doppelfolgen erfüllen, und weiter, darzutun, daß ein Teil dieser Verfahren die Permanenzbedingung auch nur für Doppelfolgen aus der genannten Klasse erfüllt[2]).

I.

Die Limitierungsverfahren, mit denen wir uns im folgenden beschäftigen, werden definiert mit Hilfe zweier Schemata:

$$\text{A.}\quad \begin{matrix} a_{00} & & & \\ a_{10} & a_{11} & & \\ \cdot & \cdot & \cdot & \\ a_{m0} & a_{m1} & \cdots & a_{mm} \\ \cdot & \cdot & \cdot & \end{matrix} \qquad \text{B.}\quad \begin{matrix} b_{00} & & & \\ b_{10} & b_{11} & & \\ \cdot & \cdot & \cdot & \\ b_{m0} & b_{m1} & \cdots & b_{mm} \\ \cdot & \cdot & \cdot & \end{matrix}$$

die den folgenden Bedingungen genügen sollen:

[1]) Diese Voraussetzung wird von Bromwich und Hardy als „condition of finitude" in ihrer Arbeit: Some extensions to multiple series of Abel's theorem on the continuity of power series, Proc. London Math. Soc. (2) 2 (1904), S. 161–189, eingeführt.

[2]) Ähnliche Untersuchungen für den Abelschen Grenzwertsatz finden sich bei H. Holzberger, Über das Verhalten von Potenzreihen mit zwei und drei Veränderlichen an der Konvergenzgrenze, Monatsh. f. Math. u. Phys. 25 (1914), S. 179–266.

a) In jeder Spalte steht eine Nullfolge, d. h. bei festem $p \geqq 0$ strebt

$$a_{mp} \to 0, \quad b_{mp} \to 0 \quad \text{für} \quad m \to \infty.$$

b) Die Zeilennormen sind beschränkt, d. h. es gibt eine Konstante K, so daß für jedes m

$$|a_{m0}| + |a_{m1}| + \ldots + |a_{mm}| < K, \quad |b_{m0}| + |b_{m1}| + \ldots + |b_{mm}| < K.$$

c) Es strebt für $m \to \infty$

$$a_{m0} + a_{m1} + \ldots + a_{mm} = A_m \to 1 \quad \text{und} \quad b_{m0} + b_{m1} + \ldots + b_{mm} = B_m \to 1.$$

Einer Doppelfolge

$$s_{\mu\nu} \qquad (\mu, \nu = 0, 1, 2, \ldots)$$

ordnen wir dann mit Hilfe dieser Schemata eine neue Doppelfolge

$$S_{mn} = \sum_{\mu,\nu=0}^{m,n} a_{m\mu} b_{n\nu} s_{\mu\nu} \qquad (m, n = 0, 1, 2, \ldots)$$

zu, und nennen die ursprüngliche Doppelfolge $s_{\mu\nu}$ $(\mu, \nu = 0, 1, 2, \ldots)$ „*$\{A, B\}$-limitierbar*“ zum Wert s, wenn

$$\lim_{m,n\to\infty} S_{mn} = s$$

ist.

Endlich bezeichnen wir eine Doppelfolge $s_{\mu\nu}$ $(\mu, \nu = 0, 1, 2, \ldots)$ als „*zur Klasse $\mathfrak{K}(A, B)$ gehörig*“, wenn die aus der Doppelfolge $s_{\mu\nu}$ $(\mu, \nu = 0, 1, 2, \ldots)$ mit Hilfe der Schemata A, B gebildete Doppelfolge $S_{m,n}$ $(m, n = 0, 1, 2, \ldots)$ beschränkt ist.

II.

1. Satz. *Eine zur Klasse $\mathfrak{K}(A, B)$ gehörige, gegen den Wert s konvergente Doppelfolge $s_{\mu\nu}$ $(\mu, \nu = 0, 1, 2, \ldots)$ ist auch $\{A, B\}$-limitierbar zum Wert s.*

Beweis. Wegen der Konvergenz der Doppelfolge $s_{\mu\nu}$ $(\mu, \nu = 0, 1, 2, \ldots)$ gegen den Wert s lassen sich zu vorgegebenem $\varepsilon > 0$ zwei Zahlen m_0, n_0 so angeben, daß

$$(1) \qquad |s_{\mu\nu} - s| < \varepsilon \quad \text{für} \quad \mu > m_0, \ \nu > n_0$$

gilt. Wir schreiben dann $S_{m,n} - s$ für $m > m_0$, $n > n_0$ in folgender Weise

$$(2) \quad S_{m,n} - s = \left\{ \sum_{\mu=0}^{m_0} \sum_{\nu=0}^{n_0} + \sum_{\mu=m_0+1}^{m} \sum_{\nu=n_0+1}^{n} + \sum_{\mu=m_0+1}^{m} \sum_{\nu=0}^{n_0} + \sum_{\mu=0}^{m_0} \sum_{\nu=n_0+1}^{n} \right\} a_{m\mu} b_{n\nu} (s_{\mu\nu} - s$$

$$+ s(A_m B_n - 1$$

$$= S_{m,n}^{(1)} + S_{m,n}^{(2)} + S_{m,n}^{(3)} + S_{m,n}^{(4)} + R_{m,n}.$$

Diese Größen schätzen wir nacheinander ab:

1. Nach Voraussetzung c) gibt es ein Zahlenpaar m_1, n_1 so, daß

$$|R_{m,n}| < \varepsilon \quad \text{für} \quad m > m_1, \quad n > n_1. \tag{3}$$

Ebenso ist für ein geeignetes Zahlenpaar m_2, n_2 $(m_2 > m_0,\ n_2 > n_0)$ nach Voraussetzung a)

$$|S^{(1)}_{m,n}| < \varepsilon \quad \text{für} \quad m > m_2, \quad n > n_2 \tag{4}$$

und endlich nach Voraussetzung b)

$$|S^{(2)}_{m,n}| < K^2\varepsilon \quad \text{für} \quad m > m_0, \quad n > n_0. \tag{5}$$

2. Eine umständlichere Untersuchung erfordern $S^{(3)}_{m,n}$ und $S^{(4)}_{m,n}$. Statt des Ausdrucks $S^{(3)}_{m,n}$ untersuchen wir

$$S^{(1)}_{m,n} + S^{(3)}_{m,n} = \sum_{\nu=0}^{n_0} \sum_{\mu=0}^{m} a_{m\mu}\, b_{n\nu}\, (s_{\mu\nu} - s).$$

Dieser Ausdruck nimmt bei Einführung der Bezeichnung

$$S^{\nu}_{m} = \sum_{\mu=0}^{m} a_{m\mu}\, (s_{\mu\nu} - s) \tag{6}$$

die Form

$$S^{(1)}_{m,n} + S^{(3)}_{m,n} = \sum_{\nu=0}^{n_0} b_{n\nu}\, S^{\nu}_{m} \tag{7}$$

an.

Mit Hilfe der Größen S^{ν}_{m} läßt sich $S_{m,n} - s$ für jedes m und n darstellen als

$$S_{m,n} - s = \sum_{\nu=0}^{n} b_{n\nu}\, S^{\nu}_{m} + R_{m,n}. \tag{8}$$

Da $s_{\mu\nu}$ $(\mu, \nu = 0, 1, 2, \ldots)$ zur Klasse $\mathfrak{K}(A, B)$ gehören soll und $R_{m,n}$ beschränkt ist, so gibt es nach (8) eine Konstante C, mit der

$$\left| \sum_{\nu=0}^{n} b_{n\nu}\, S^{\nu}_{m} \right| < C \qquad (m, n = 0, 1, 2, \ldots) \tag{9}$$

gilt. Andrerseits lassen sich die Größen S^{ν}_{m} mit $\nu > n_0$ wie folgt abschätzen:

$$|S^{\nu}_{m}| \leqq \left| \sum_{\mu=0}^{m_0} a_{m\mu}\, (s_{\mu\nu} - s) \right| + \left| \sum_{\mu=m_0+1}^{m} a_{m\mu}\, (s_{\mu\nu} - s) \right| \leqq K \max_{0 \leqq \mu \leqq m_0} |s_{\mu\nu} - s| + K\varepsilon,$$

da ja für $\mu > m_0$, $\nu > n_0$ die Abschätzung (1) zuständig ist. Der letzte Teil der vorstehenden Ungleichungen hängt außer von der fest gewählten Zahl m_0 nur von ν ab, wir bezeichnen ihn daher mit c_ν, also

$$|S^{\nu}_{m}| < c_\nu \quad \text{für} \quad \nu > n_0, \quad m = 0, 1, 2, \ldots .$$

Daraus folgt für $n > n_0$

$$\left| \sum_{\nu=n_0+1}^{n} b_{n\nu}\, S^{\nu}_{m} \right| < K \operatorname{Max}(c_{n_0+1}, \ldots, c_n) = K\, c'_n, \tag{10}$$

und aus (9) und (10) zusammen schließt man. daß es zu jedem n eine Zahl $C^{(n)}$ gibt, mit der (vgl. Fußnote [3]))

$$\left| \sum_{\nu=0}^{n_0} b_{n\nu} S_m^{\nu} \right| < C^{(n)} \qquad \begin{pmatrix} m = 0, 1, 2, \ldots \\ n = 0, 1, 2, \ldots \end{pmatrix}$$

gilt. Für jedes m kann man also setzen

$$\begin{aligned}
&b_{00} \quad S_m^0 && = D_m^{(0)} \\
&b_{10} \quad S_m^0 + b_{11} \quad S_m^1 && = D_m^{(1)} \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&b_{n_0 0} \quad S_m^0 + b_{n_0 1} \quad S_m^1 + \ldots + b_{n_0 n_0} \quad S_m^{n_0} && = D_m^{(n_0)} \\
&b_{n_0+1,0} \, S_m^0 + b_{n_0+1,1} \, S_m^1 + \ldots + b_{n_0+1,n_0} \, S_m^{n_0} && = D_m^{(n_0+1)} \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots
\end{aligned} \tag{11}$$

und dabei ist

$$|D_m^{(\varrho)}| < C^{(\varrho)} \qquad \begin{pmatrix} m = 0, 1, 2, \ldots \\ \varrho = 0, 1, 2, \ldots \end{pmatrix}. \tag{12}$$

Nun betrachten wir das unendliche Schema der Größen

$$\begin{aligned}
&b_{00} \quad 0 \quad \ldots 0 \\
&b_{10} \quad b_{11} \quad \ldots 0 \\
&\ldots\ldots\ldots\ldots \\
&b_{n_0 0} \quad b_{n_0 1} \quad \ldots b_{n_0 n_0} \\
&b_{n_0+1,0} \, b_{n_0+1,1} \ldots b_{n_0+1,n_0} \\
&\ldots\ldots\ldots\ldots
\end{aligned} \tag{13}$$

Sind alle diese Größen Null, so ist auch

$$S_{m,n}^{(1)} + S_{m,n}^{(3)} = 0. \tag{14}$$

Andernfalls gibt es eine Zahl k mit $0 < k \leqq n_0 + 1$, so daß sich aus dem Schema (13) mindestens eine nicht verschwindende k-reihige Determinante [3])

$$\begin{vmatrix} b_{\sigma_1 \tau_1} & b_{\sigma_1 \tau_2} & \ldots & b_{\sigma_1 \tau_k} \\ b_{\sigma_2 \tau_1} & b_{\sigma_2 \tau_2} & \ldots & b_{\sigma_2 \tau_k} \\ \ldots & \ldots & \ldots & \ldots \\ b_{\sigma_k \tau_1} & b_{\sigma_k \tau_2} & \ldots & b_{\sigma_k \tau_k} \end{vmatrix} \qquad \begin{pmatrix} 0 \leqq \tau_1 < \tau_2 < \ldots < \tau_k \leqq n_0 \\ 0 \leqq \sigma_1 < \sigma_2 < \ldots < \sigma_k \end{pmatrix} \tag{15}$$

bilden läßt, während alle $(k+1)$-reihigen Determinanten verschwinden. Dann existiert aber für jedes $n \geqq 0$ ein System von Größen

$$\lambda_1^{(n)}, \lambda_2^{(n)}, \ldots, \lambda_k^{(n)},$$

---

[3]) Größen $b_{\sigma_i \tau_\varkappa}$ mit $\tau_\varkappa > \sigma_i$ sind gleich Null zu setzen.

das die Gleichungen

$$\begin{aligned} b_{\sigma_1 0}\,\lambda_1^{(n)} + b_{\sigma_2 0}\,\lambda_2^{(n)} + \ldots + b_{\sigma_k 0}\,\lambda_k^{(n)} &= b_{n,0} \\ b_{\sigma_1 1}\,\lambda_1^{(n)} + b_{\sigma_2 1}\,\lambda_2^{(n)} + \ldots + b_{\sigma_k 1}\,\lambda_k^{(n)} &= b_{n,1} \\ \cdot\ \cdot\ \cdot

$\sigma_0 > P$, *so daß die Determinanten*

$$\begin{vmatrix} a_{\varrho_0,0} & a_{\varrho_0,1} & \dots & a_{\varrho_0,p} \\ a_{\varrho_1,0} & a_{\varrho_1,1} & \dots & a_{\varrho_1,p} \\ \dots & \dots & \dots & \dots \\ a_{\varrho_p,0} & a_{\varrho_p,1} & \dots & a_{\varrho_p,p} \end{vmatrix} \quad \text{und} \quad \begin{vmatrix} b_{\sigma_0,0} & b_{\sigma_0,1} & \dots & b_{\sigma_0,p} \\ b_{\sigma_1,0} & b_{\sigma_1,1} & \dots & b_{\sigma_1,p} \\ \dots & \dots & \dots & \dots \\ b_{\sigma_p,0} & b_{\sigma_p,1} & \dots & b_{\sigma_p,p} \end{vmatrix}$$

von Null verschieden sind.

Dann gehört jede zugleich konvergente und $\{A, B\}$*-limitierbare Doppelfolge* $s_{\mu\nu}$ $(\mu, \nu = 0, 1, 2, \dots)$ *zur Klasse* $\mathfrak{K}(A, B)$.

Beweis: 1. Die Doppelfolge $s_{\mu\nu}$ $(\mu, \nu = 0, 1, 2, \dots)$ möge gegen s, die transformierte Doppelfolge $S_{m,n}$ $(m, n = 0, 1, 2, \dots)$ gegen S streben. Dann läßt sich zu vorgegebenem $\varepsilon > 0$ ein Zahlenpaar m_0, n_0 so bestimmen, daß

$$|s_{\mu\nu} - s| < \varepsilon \quad \text{für} \quad \mu > m_0, \ \nu > n_0 \tag{20}$$

und

$$|S_{m,n} - S| < \varepsilon \quad \text{für} \quad m > m_0, \ n > n_0 \tag{21}$$

gilt. Wir betrachten nun wieder den Ausdruck $S_{m,n} - s$ und zerlegen ihn in derselben Weise wie in (2). Man sieht dann unmittelbar ein, daß es ein Zahlenpaar $M, N (M > m_0, N > n_0)$ gibt, mit dem

$$|R_{m,n}| < \varepsilon, \ |S^{(1)}_{m,n}| < \varepsilon, \ |S^{(2)}_{m,n}| < K^2\varepsilon \quad \text{für} \quad m > M, n > N \tag{22}$$

gilt. Aus (21) und (22) folgt, daß für eine geeignete Konstante C_1

$$|S^{(3)}_{m,n} + S^{(4)}_{m,n}| < C_1 \quad \text{für} \quad m > M, n > N$$

also zufolge der zweiten Relation (22) für eine passende Konstante C_2 auch

$$|(S^{(1)}_{m,n} + S^{(3)}_{m,n}) + (S^{(1)}_{m,n} + S^{(4)}_{m,n})| < C_2 \quad \text{für} \quad m > M, n > N \tag{23}$$

gelten muß. Daraus werden wir folgern, daß die sämtlichen Größen $S_{m,n}$ mit $n \leqq n_0$ beschränkt sein müssen. Genau ebenso kann man die Beschränktheit der Größen $S_{m,n}$ mit $m \leqq m_0$ nachweisen, und aus diesen Tatsachen ergibt sich unter Berücksichtigung von (21) die Behauptung.

2. Um die Beschränktheit der Größen $S_{m,n}$ mit $n \leqq n_0$ nachzuweisen[4]), wählen wir zu n_0 und N ein nach Voraussetzung d) vorhandenes System von Indizes $\sigma_0, \sigma_1, \dots, \sigma_{n_0}$, für das die zweite der in d) genannten Determinanten nicht verschwindet $(N < \sigma_0 < \sigma_1 < \dots < \sigma_{n_0})$. Dann gibt es gewiß eine Konstante C_3 so, daß

$$\left|\sum_{\nu=0}^{n} b_{n\nu}(s_{\mu\nu} - s)\right| < C_3 \quad \text{für} \quad \begin{matrix} \mu = 0, 1, 2, \dots, m_0 \\ n = \sigma_0, \sigma_1, \sigma_2, \dots, \sigma_{n_0} \end{matrix}$$

---

[4]) Einer ähnlichen Schlußweise bedient sich Herr Hartogs; vgl. F. Hartogs, Beiträge zur elementaren Theorie der Potenzreihen und der eindeutigen analytischen Funktionen zweier Veränderlichen, Dissertation München 1904, S. 21.

und daher nach Voraussetzung b)

$$\left|\sum_{\mu=0}^{m_0}\sum_{\nu=0}^{n} a_{m\mu}\, b_{n\nu}(s_{\mu\nu}-s)\right| < K C_3 \quad \text{für} \quad \begin{matrix} m = 0, 1, 2, \ldots \\ n = \sigma_0, \sigma_1, \ldots, \sigma_{n_0} \end{matrix}$$

gilt. Vermöge (23) muß also auch

$$\left|\sum_{\mu=0}^{m}\sum_{\nu=0}^{n_0} a_{m\mu}\, b_{n\nu}(s_{\mu\nu}-s)\right| < C_2 + K C_3 = C_4 \quad \text{für} \quad \begin{matrix} m > M \\ n = \sigma_0, \sigma_1, \ldots, \sigma_{n_0} \end{matrix}$$

oder unter Verwendung der Bezeichnung (6)

$$\left|\sum_{\nu=0}^{n_0} b_{n\nu} S_m^{\nu}\right| < C_4 \quad \text{für} \quad \begin{matrix} m > M \\ n = \sigma_0, \sigma_1, \ldots, \sigma_{n_0} \end{matrix}$$

sein, d. h. für die Größen

$$\begin{aligned} &b_{\sigma_0 0} S_m^0 + b_{\sigma_0 1} S_m^1 + \ldots + b_{\sigma_0 n_0} S_m^{n_0} = E_m^{(0)} \\ &b_{\sigma_1 0} S_m^0 + b_{\sigma_1 1} S_m^1 + \ldots + b_{\sigma_1 n_0} S_m^{n_0} = E_m^{(1)} \\ &\quad\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &b_{\sigma_{n_0} 0} S_m^0 + b_{\sigma_{n_0} 1} S_m^1 + \ldots + b_{\sigma_{n_0} n_0} S_m^{n_0} = E_m^{(n_0)} \end{aligned} \tag{24}$$

gilt

$$|E_m^{(i)}| < C_4 \quad \text{für} \quad \begin{matrix} i = 0, 1, \ldots, n_0 \\ m > M \end{matrix}. \tag{25}$$

Da die Determinante des Systems (24) nicht verschwindet, so schließt man, daß es eine von m unabhängige Konstante C (die nur von den fest gewählten Größen m_0, n_0, M, $b_{\sigma\kappa,\lambda}$ abhängt) gibt, mit der

$$|S_m^0| < C, \ |S_m^1| < C, \ldots, |S_m^{n_0}| < C \quad \text{für} \quad m > M \tag{26}$$

gilt.

Wir wählen nun einen Index $n^* \leqq n_0$; dann wird

$$S_{m,n^*} - s = b_{n^*,0} S_m^0 + b_{n^*,1} S_m^1 + \ldots + b_{n^*,n^*} S_m^{n^*} + R_{m,n^*}$$

und daraus ergibt sich mit Hilfe der Voraussetzungen b) und c) sowie der Abschätzung (26) die Beschränktheit von S_{m,n^*} für $m = 0, 1, 2, \ldots$

Bemerkung: Aus Satz 1 und 2 zusammen ergibt sich, daß für eine gegen den Wert s konvergierende Doppelfolge $s_{\mu\nu}$ $(\mu, \nu = 0, 1, 2, \ldots)$, die zugleich $\{A, B\}$-limitierbar zum Wert S ist, $S = s$ sein muß, falls das Limitierungsverfahren den Voraussetzungen a) bis d) genügt. Wie wir im folgenden zeigen wollen, läßt sich diese Tatsache $(S = s)$ aber auch direkt und ohne Zuhilfenahme von d) beweisen, so daß sie also für alle von uns betrachteten Limitierungsverfahren gilt.

Zum Beweis ist offenbar nur zu zeigen, daß es zu vorgegebenem $\varepsilon > 0$ Größen $S_{m,n}$, deren beide Indizes nach ∞ streben, gibt, für die $|S_{m,n} - s| < \varepsilon$ ist, denn dann kann die Folge $S_{m,n}$ $(m, n = 0, 1, \ldots)$ gewiß gegen keinen

von s verschiedenen Grenzwert konvergieren. Wir knüpfen dazu an den ersten Teil des Beweises von Satz 2 an und verfahren weiterhin ähnlich wie beim Beweis von Satz 1. In dem unendlichen Schema der Größen

$$\begin{matrix} b_{N+1,0}\, b_{N+1,1} \dots b_{N+1,n_0} \\ b_{N+2,0}\, b_{N+2,1} \dots b_{N+2,n_0} \\ \dots\dots\dots\dots \end{matrix} \tag{27}$$

sind entweder

a) alle Größen Null, dann ist

$$\sum_{\nu=0}^{n_0} b_{n\nu}\, S_m^\nu = 0 \quad \text{für} \quad n > N \tag{28}$$

oder aber

b) es gibt eine Zahl k mit $0 < k < 1$, so daß die k-reihige Determinante

$$\begin{vmatrix} b_{\alpha_1,\beta_1}\, b_{\alpha_1,\beta_2} \dots b_{\alpha_1,\beta_k} \\ b_{\alpha_2,\beta_1}\, b_{\alpha_2,\beta_2} \dots b_{\alpha_2,\beta_k} \\ \dots\dots\dots\dots \\ b_{\alpha_k,\beta_1}\, b_{\alpha_k,\beta_2} \dots b_{\alpha_k,\beta_k} \end{vmatrix} \qquad \begin{pmatrix} N < \alpha_1 < \alpha_2 < \dots < \alpha_k \\ 0 \leqq \beta_1 < \beta_2 < \dots < \beta_k \leqq n_0 \end{pmatrix} \tag{29}$$

nicht verschwindet, während alle $(k+1)$-reihigen Determinanten aus dem Schema (27) verschwinden. Setzt man dann allgemein

$$b_{n,0}\, S_m^0 + b_{n,1}\, S_m^1 + \dots + b_{n,n_0} S_m^{n_0} = F_m^{(n)}, \quad (n = 0, 1, 2, \dots), \tag{30}$$

so erkennt man genau wie beim Beweis von Satz 1, daß es zu jedem $n > N$ ein System von Größen

$$\lambda_1^{(n)}, \lambda_2^{(n)}, \dots, \lambda_k^{(n)}$$

gibt, für das

$$F_m^{(n)} = \lambda_1^{(n)} F_m^{(\alpha_1)} + \lambda_2^{(n)} F_m^{(\alpha_2)} + \dots + \lambda_k^{(n)} F_m^{(\alpha_k)} \qquad \begin{pmatrix} m = 0, 1, 2, \dots \\ n > N \end{pmatrix} \tag{31}$$

und

$$|\lambda_i^{(n)}| < \varepsilon \quad \text{für} \quad \begin{matrix} i = 0, 1, \dots, k \\ n > N' > N \end{matrix} \tag{32}$$

ist. Weiter zeigt man wie im 2. Teil des Beweises von Satz 2 (vgl. die auf (24) und (25) führende Abschätzung, wo nur $\sigma_0, \sigma_1, \dots, \sigma_{n_0}$ durch $\alpha_1, \alpha_2, \dots, \alpha_k$ zu ersetzen ist), daß

$$|F_m^{(\alpha_i)}| < C_5 \quad \text{für} \quad \begin{matrix} i = 0, 1, \dots, k \\ m > M \end{matrix} \tag{33}$$

gilt, wo C_5 eine passend gewählte, von m unabhängige Konstante bedeutet.

Wir betrachten nun die Größen

$$S_m^\nu = \sum_{\mu=0}^{m} a_{m\mu}(s_{\mu\nu} - s) \quad \text{für} \quad \nu > n_0.$$

Nach (20) ist

$$|S_m^\nu| \leqq \Big|\sum_{\mu=0}^{m_0} a_{m\mu}(s_{\mu\nu}-s)\Big| + \Big|\sum_{\mu=m_0+1}^{m} a_{m\mu}(s_{\mu\nu}-s)\Big|$$

$$\leqq \operatorname*{Max}_{\mu=0,1,\ldots,m_0} |s_{\mu\nu}-s| \sum_{\mu=0}^{m_0} a_{m\mu} + K\varepsilon,$$

und daraus folgt nach Voraussetzung a)

$$(34) \qquad |S_m^\nu| < 2K\varepsilon \quad \text{für} \quad \begin{matrix} \nu > n_0 \\ m > m(\nu) \end{matrix}.$$

Um den Beweis vollends zu Ende zu führen, wählen wir einen beliebigen Index $n = \bar{n} > N'$ und betrachten $S_{m,\bar{n}} - s$ für $m > M'$, wo $M' = \operatorname{Max}\{M, m(n_0+1), m(n_0+2), \ldots, m(\bar{n})\}$ sein soll. Nach (30) ist:

$$S_{m,\bar{n}} - s = F_m^{(\bar{n})} + \sum_{\nu=n_0+1}^{\bar{n}} b_{\bar{n},\nu} S_m^\nu + R_{m,\bar{n}}.$$

Zufolge (28) bzw. (32) und (33) ist

$$|F_m^{(\bar{n})}| < \varepsilon C_5 \quad \text{für} \quad m > M.$$

Andererseits gilt nach (34):

$$\Big|\sum_{\nu=n_0+1}^{\bar{n}} b_{\bar{n},\nu} S_m^\nu\Big| \leqq \sum_{\nu=n_0+1}^{n} |b_{\bar{n},\nu}|\,|S_m^\nu| \leqq 2K^2\varepsilon \quad \text{für} \quad m > M'.$$

Zusammenfassend ergibt sich also, wenn man noch (22) heranzieht,

$$|S_{m,\bar{n}} - s| < \varepsilon(1 + C_5 + 2K^2) \quad \text{für} \quad m > M',$$

und da $\bar{n}$ willkürlich, nur hinreichend groß gewählt ist, ist der Beweis fertig.

IV.

Die Limitierungsverfahren $\{A, B\}$, welche den sämtlichen Voraussetzungen a) bis d) genügen, haben folgende bemerkenswerte Eigenschaft: Ist eine gegen den Wert s konvergente Doppelfolge $s_{\mu\nu}$ $(\mu, \nu = 0, 1, 2, \ldots)$ zugleich $\{A, B\}$-limitierbar zum Wert s, so gehört $s_{\mu\nu}$ $(\mu, \nu = 0, 1, \ldots)$ nach Satz 2 zur Klasse $\mathfrak{K}(A, B)$, aber auch die transformierte Folge $S_{m,n}$ $(m, n = 0, 1, 2, \ldots)$ gehört zur Klasse $\mathfrak{K}(A, B)$ (sie genügt sogar der „condition of finitude"), ist also selbst wieder $\{A, B\}$-limitierbar zum Wert s usw.

Als Beispiel eines derartigen Verfahrens erwähnen wir das E_1-Verfahren, bei dem die Folge $S_{m,n}$ definiert ist vermöge

$$S_{m,n} = \frac{1}{2^{m+n}} \sum_{\mu=0}^{m} \sum_{\nu=0}^{n} \binom{m}{\mu} \binom{n}{\nu} s_{\mu\nu} \qquad (m, n = 0, 1, 2, \ldots).$$

Hier sind die Voraussetzungen a) bis c) offensichtlich erfüllt, aber auch d), denn wählt man $q > \mathrm{Max}(P, p)$, so ist die Determinante

$$\Delta(q,p) = \begin{vmatrix} \binom{q}{0} & \binom{q}{1} & \cdots & \binom{q}{p} \\ \binom{q+1}{0} & \binom{q+1}{1} & \cdots & \binom{q+1}{p} \\ \cdot & \cdot & \cdots & \cdot \\ \binom{q+p}{0} & \binom{q+p}{1} & \cdots & \binom{q+p}{p} \end{vmatrix} = 1,$$

wie man mittels der leicht nachzuprüfenden Relation

$$\Delta(q,p) = \binom{q}{0} \Delta(q, p-1)$$

erkennt. Ist also eine gegen den Wert s konvergente Doppelfolge zugleich E_1-limitierbar, so ist sie für jedes ganzzahlige r auch E_r-limitierbar zum Wert s.

Das H_1-Verfahren, bei dem

$$S_{m,n} = \frac{1}{(m+1)(n+1)} \sum_{\mu=0}^{m} \sum_{\nu=0}^{n} s_{m,n} \quad (m, n = 0, 1, 2, \ldots)$$

ist, genügt zwar den Voraussetzungen a), b), c) — es ist also Satz 1 anwendbar — nicht aber d), wie man sofort daraus ersieht, daß alle $a_{m\mu}$ für $\mu = 0, 1, \ldots, m$ gleich sind, die fraglichen Determinanten also in jeder Zeile lauter gleiche Elemente aufweisen. In der Tat folgt aus der H_1-Limitierbarkeit einer konvergenten Doppelfolge $s_{\mu\nu}$ $(\mu, \nu = 0, 1, 2, \ldots)$ weder die Beschränktheit noch die H_1-Limitierbarkeit der Doppelfolge $S_{m,n}$ $(m, n = 0, 1, 2, \ldots)$, wie das Beispiel

$$(s_{\mu\nu}) = \begin{pmatrix} 0 & 2 & 4 & 6 & \ldots \\ 0 & -2 & -4 & -6 & \ldots \\ 0 & 0 & 0 & 0 & \ldots \\ 0 & 0 & 0 & 0 & \ldots \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}, \quad \text{also} \quad (S_{m,n}) = \begin{pmatrix} 0 & 1 & 2 & 3 & \ldots \\ 0 & 0 & 0 & 0 & \ldots \\ 0 & 0 & 0 & 0 & \ldots \\ 0 & 0 & 0 & 0 & \ldots \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

zeigt.

Stuttgart, 18. 2. 1931.

(Eingegangen am 25. Februar 1931.)

Über das Verhalten der Potenzreihen auf dem Rande des Konvergenzkreises II.

Von

Friedrich Lösch in Stuttgart.

Die von mir kürzlich unter gleichem Titel[1]) angegebenen Beispiele zum Verhalten der Potenzreihen auf dem Rande des Konvergenzkreises möchte ich im folgenden ergänzen durch ein sehr einfaches Beispiel einer Potenzreihe, deren Konvergenz- und Divergenzpunkte auf dem Rande des Konvergenzkreises je überall dicht liegen[2]).

Die Potenzreihe

$$\sum_{n=0}^{\infty} a_n x^n, \tag{1}$$

die aus der Polynomreihe

$$\sum_{n=1}^{\infty} P_n(x) = \sum_{n=1}^{\infty} \frac{1}{n}\left[\frac{1}{2}x(1+x^{2^n})\right]^{2^{2^n}} \tag{2}$$

entsteht, wenn die Polynome in ihre einzelnen Glieder zerlegt werden, konvergiert und divergiert je in einer überall dichten Punktmenge des Einheitskreises.

Zum Beweis zeigen wir:

1. Es genügt schon, die Behauptung für die Polynomreihe (2) nachzuweisen: Die 2^{2^ν}-te bis $(2^{2^{\nu+1}}-1)$-te Partialsumme $(s_\nu = \sum_{n=0}^{\nu} a_n x^n)$ der Reihe (1) unterscheidet sich nämlich von der ν-ten Partialsumme $(S_\nu = \sum_{n=1}^{\nu} P_n(x))$ der

[1]) F. Lösch, Über das Verhalten der Potenzreihen auf dem Rande des Konvergenzkreises, Math. Zeitschrift **33** (1931), S. 791—795.

[2]) Andere Beispiele solcher Potenzreihen finden sich z. B. bei L. Neder, Konvergenzdefekte der Potenzreihen stetiger Funktionen auf dem Rande des Konvergenzkreises, Math. Zeitschrift **6** (1920), S. 262—269.

Reihe (2) auf $|x|=1$ höchstens um $\frac{1}{\nu}$, da jede Summe von Gliedern der Binomialentwicklung von

$$\left[\frac{1}{2}x(1+x^{2^\nu})\right]^{2^{2^\nu}}$$

auf $|x|=1$ höchstens den Wert 1 hat.

2. Die Reihe (2) divergiert für die sämtlichen 2^m-ten Einheitswurzeln, wobei m jeden positiven, ganzzahligen Wert haben kann: Ist nämlich $\zeta^{2^m}=1$, so ist

$$\left[\frac{1}{2}\zeta(1+\zeta^{2^n})\right]^{2^{2^n}}=1 \quad \text{für } n \geqq m.$$

3. Die Reihe (2) konvergiert für jede von $+1$ verschiedene 3^m-te Einheitswurzel, wobei m wieder jeden ganzen positiven Wert haben kann: Ist nämlich $\zeta^{3^m}=1, \zeta \neq 1$ und bezeichnet man die von $+1$ verschiedenen 3^m-ten Einheitswurzeln mit $\zeta_1, \zeta_2, \ldots, \zeta_{3^m-1}$, so ist ζ^{2^n} für jedes n gleich einem ζ_ϱ $(\varrho=1, 2, \ldots, 3^m-1)$, also

$$|1+\zeta^{2^n}| \leqq \text{Max}\{|1+\zeta_1|, |1+\zeta_2|, \ldots, |1+\zeta_{3^m-1}|\} \leqq q < 2,$$

und somit

$$\left|\left[\frac{1}{2}\zeta(1+\zeta^{2^n})\right]^{2^{2^n}}\right| < \left(\frac{q}{2}\right)^{2^{2^n}}.$$

Bemerkung: Das Beispiel ist vom gleichen Typus wie das in der eingangs genannten Note gegebene Beispiel einer Potenzreihe, die auf $|x|=1$ überall, aber nur bedingt konvergiert und die einfach durch Zerlegung der Polynome der Reihe

$$\sum_{n=1}^{\infty}\frac{(-1)^n}{n}\left[\frac{1}{2}x(1+x)\right]^{4^n}$$

in ihre einzelnen Glieder entsteht. (Diese Eigenschaft ergibt sich sofort daraus, daß $\frac{1}{2}x(1+x)=1$ für $x=1$, $|\frac{1}{2}x(1+x)|<1$ für $|x|=1, x \neq 1$ ist.)

Zwischen diesen beiden Beispielen steht gewissermaßen noch die Potenzreihe, die dadurch entsteht, daß die Polynome der Reihe

$$\sum_{n=1}^{\infty}\frac{1}{n}\left[\frac{1}{2}x(1+x)\right]^{4^n}$$

in ihre einzelnen Glieder zerlegt werden und von der man unmittelbar erkennt, daß sie in allen von $x=1$ verschiedenen Punkten des Einheitskreises konvergiert, dagegen in $x=1$ divergiert.

Stuttgart, 21. 2. 1931.

(Eingegangen am 25. Februar 1931.)

Das Eigenwertproblem der selbstadjungierten linearen Differentialgleichung vierter Ordnung*).

Von

Hermann Boerner in Leipzig.

Herr Lichtenstein hat in mehreren Abhandlungen zur Erledigung der Eigenwertprobleme linearer Differentialgleichungen zweiter Ordnung die von Hilbert begründete Methode der unendlichvielen Veränderlichen herangezogen. Seiner Abhandlung über die gewöhnliche („Sturm-Liouvillesche") Differentialgleichung zweiter Ordnung[1]) folgt die vorliegende Arbeit, in der die entsprechenden Sätze für die selbstadjungierte lineare Differentialgleichung vierter Ordnung

$$(A) \quad \frac{d^2}{dx^2}\left(p(x)\frac{d^2u}{dx^2}\right) + \frac{d}{dx}\left([\lambda r(x) + r_1(x)]\frac{du}{dx}\right) - [\lambda k(x) + k_1(x)]u(x) = 0$$
$$(p > 0)$$

hergeleitet werden, im wesentlichen unter Beschränkung auf das einfachste Randwertproblem, dem — in den einfachsten Fällen — in der Physik das Problem der Schwingungen eines beiderseits eingespannten Stabes entspricht. Dieses Problem, zuerst in der „Theory of Sound" des Lord Rayleigh ausführlich behandelt, wurde von M. A. Davidoglou[2]) wieder aufgenommen, der für die Gleichung

$$\frac{d^2}{dx^2}\left(p\frac{d^2u}{dx^2}\right) - \lambda k u = 0$$

*) Erster Teil der von der philosophischen Fakultät der Universität Leipzig als Inaugural-Dissertation zugelassenen Arbeit „Über einige Eigenwertprobleme und ihre Anwendung in der Variationsrechnung".

[1]) L. Lichtenstein, Zur Analysis der unendlichen vielen Variablen. Erste Abhandlung. Rend. circ. mat. Palermo **38** (1914), S. 113–166. Anwendung seiner Methode auf Systeme von Differentialgleichungen zweiter Ordnung durch H. Geiringer, Über eine Randwertaufgabe usw., Math. Zeitschr. **12** (1922), S. 1–17.

[2]) M. A. Davidoglou, Sur l'équation des vibrations transversales etc., Ann. Ec. Norm. 1900, S. 359–444, und Sur l'équation etc., Ann. Ec. Norm. 1905, S. 539–565.

unter der Voraussetzung $k > 0$ die Eigenwerte und Eigenfunktionen herleitet. Entwicklungssätze wurden von J. Tamarkine und W. Stekloff[3]) aufgestellt. Sie beweisen für die einfachste Gleichung den Entwicklungssatz im gleichen Umfange, wie er für die Fouriersche Entwickelung gilt, allerdings nur für das Innere des betrachteten Intervalls. Der von uns benutzten Methode sind nur Funktionen zugänglich, die mindestens stetig sind, stetige Ableitung erster Ordnung besitzen und der Randbedingung genügen. Für solche Funktionen werden wir die gleichmäßige Konvergenz der (gliedweise differenzierbaren) Entwicklung im abgeschlossenen Intervall beweisen. Allgemeine „Sturmsche Systeme" von Randbedingungen behandelt S. A. Janczewsky[4]) für die Gleichung

$$\frac{d^2}{dx^2}\left(p\,\frac{d^2u}{dx^2}\right) - (\lambda k + k_1)\,u = 0$$

unter der Voraussetzung $k \geqq 0$. *Wir werden über das Vorzeichen von $r(x)$ und $k(x)$ keinerlei Voraussetzung machen.* Kürzlich hat G. Cimmino[5]) die Existenzsätze für die Eigenwerte und Eigenfunktionen derselben allgemeinen Differentialgleichung (A) bewiesen, die in der gegenwärtigen Arbeit behandelt wird. Er setzt $r_1 \leqq 0$, $k_1 \leqq 0$ voraus, — was wir durch eine etwas allgemeinere, aber undurchsichtigere Voraussetung ersetzen können —, und schließt den Fall aus, daß r und k im ganzen Intervall das entgegengesetzte Vorzeichen haben. Diese Voraussetzung werden wir nötig haben, um gewisse Aussagen über das Vorzeichen der Eigenwerte machen zu können; *zum Beweise der Existenz der Eigenwerte und Eigenfunktionen und für die Entwicklungssätze brauchen wir sie nicht.* Ein wesentlicher Vorteil unserer Methode ist ferner, daß die Existenz- und Entwicklungssätze, einmal für die Differentialgleichung vierter Ordnung bewiesen, ganz unmittelbar auf die Differentialgleichung von der Ordnung $2n$ übertragen werden können. Zuletzt wird eine Integro-Differentialgleichung vierter Ordnung auf dieselbe Art behandelt.

In einer demnächst erscheinenden Arbeit sollen dann einige Ergebnisse der vorliegenden auf die Theorie der zweiten Variation bei Variationsproblemen von der Form

$$\delta \int_a^b f(x, y, y', y'')\,dy = 0$$

---

[3]) W. Stekloff et J. Tamarkine, Problème des vibrations transversales etc., Rend. circ. mat. Palermo **31** (1911), S. 341—362.

[4]) S. A. Janczewsky, Oscillation theorems for the differential boundary value problems etc., Ann. of Math. 1928, S. 521—542. Vgl. auch O. Haupt, Untersuchungen über Oszillationstheoreme, Dissertation, Würzburg 1911, Zweiter Teil.

[5]) G. Cimmino, Autosoluzioni e autovalori nelle equazioni etc., Math. Zeitschr. **32** (1930), S. 4—58.

angewendet werden. Um dort auch den Fall betrachten zu können, daß die Endpunkte der Extremalen auf zwei vertikalen Geraden beweglich sind oder daß die Richtung der Extremalen in diesen Punkten variabel ist, werden im siebenten und achten Abschnitt der gegenwärtigen Arbeit zwei spezielle Randwertaufgaben behandelt, bei denen der Parameter in den Randbedingungen vorkommt. Endlich wird uns die zuletzt betrachtete Integro-Differentialgleichung in den Stand setzen, auch über das zugehörige isoperimetrische Problem einige Sätze herzuleiten. Die Grundlage werden bei alledem wiederum einige Abhandlungen von Herrn Professor Lichtenstein bilden, dem ich auch für die Anregung zu diesen Arbeiten und für seine Anteilnahme an ihrem Fortgang zu großem Dank verpflichtet bin.

---

1. Es sei $p(x)$ eine gegebene samt ihren Ableitungen erster und zweiter Ordnung im Intervall $\langle a\,b\rangle$ stetige positive Funktion, die nirgends verschwindet, und es seien $\omega_i(x)$ $(i = 1, 2, \ldots)$ die den Bedingungen

$$\int_a^b \omega_i(x)\,\omega_j(x)\,dx = \delta_{ij}$$

gemäß normierten Eigenfunktionen, ν_i die zugehörigen Eigenwerte der Differentialgleichung

$$\frac{d^2}{dx^2}\left(p(x)\frac{d^2u}{dx^2}\right) - \nu\,u(x) = 0, \tag{1}$$

die zur Randbedingung

$$u(a) = u'(a) = 0, \qquad u(b) = u'(b) = 0 \tag{2}$$

gehören. Eine Funktion $u(x)$ soll eine „*zulässige Funktion*“ heißen, wenn sie samt ihrer ersten Ableitung in $\langle a\,b\rangle$ stetig ist, der Randbedingung (2) genügt und wenn über ihre zweite Ableitung noch eine Voraussetzung erfüllt ist, die weiter unten präzisiert werden soll[6]). Wir zeigen, daß sich jede zulässige Funktion in eine nach den Funktionen $\omega_i(x)$ fortschreitende unbedingt und gleichmäßig konvergente Reihe entwickeln läßt.

Es sei $G(\xi, x)$ die Greensche Funktion der Differentialgleichung

$$\frac{d^2}{dx^2}\left(p(x)\frac{d^2u}{dx^2}\right) = 0 \tag{3}$$

zur Randbedingung (2), die folgendermaßen gewonnen werden kann. Sind

---

[6]) Sie läuft darauf hinaus, daß im Lebesgueschen Sinne $\int_a^b p\left(\frac{d^2u}{dx^2}\right)^2 dx$ existiert und $\int_u^x \frac{d^2u}{dx^2}\,dx = \frac{du}{dx}$ ist. Doch wird es sich nachher erübrigen, mit diesem Begriff zu operieren.

$u_1(x)$ und $u_2(x)$ zwei linear unabhängige Lösungen von (3), die beide für $x = a$ samt ihrer Ableitung verschwinden, $u_3(x)$ und $u_4(x)$ zwei Lösungen, für die dasselbe im Punkte b gilt, so bilden die Funktionen $u_1, \ldots, u_4$ ein Fundamentalsystem linear unabhängiger Lösungen, und die zugehörige Wronskische Determinante ist überall von Null verschieden. Setzen wir darum

$$G(\xi, x) = \begin{cases} c_1 u_1 + c_2 u_2 & (x < \xi), \\ c_3 u_3 + c_4 u_4 & (x > \xi), \end{cases}$$

so werden die Konstanten $c_1, \ldots, c_4$ eindeutig durch die Forderung bestimmt, daß für $x = \xi$ — die partiellen Ableitungen seien durch Indizes bezeichnet — $G(\xi, x)$, $G_x(\xi, x)$, $G_{xx}(\xi, x)$ als Funktionen von x stetig sein und die dritte Ableitung den Sprung

$$\frac{\partial}{\partial x}(p\, G_{xx}(\xi, x))\Big|_{x=\xi-0}^{x=\xi+0} = -1$$

aufweisen soll. Man erhält Ausdrücke für $c_1, \ldots, c_4$ als Funktionen von ξ. Differenziert man diese nach ξ, so verifiziert man leicht, daß die Funktion $G_\xi(\xi, x)$ in genau derselben Weise wie $G(\xi, x)$ gewonnen werden kann, wenn man die Stetigkeit von $G_\xi(\xi, x)$, $G_{\xi x}(\xi, x)$, $\frac{\partial}{\partial x}(p\, G_{\xi xx}(\xi, x))$ und den Sprung

$$p\, G_{\xi xx}(\xi, x)\Big|_{x=\xi-0}^{x=\xi+0} = 1$$

fordert. *Demnach sind* $G(\xi, x)$ *und* $G_\xi(\xi, x)$ *„zulässige Funktionen“.* Bezeichnet $u(x)$ eine beliebige zulässige Funktion, so gelten, wie man durch teilweise Integration verifiziert, die Relationen

$$\int_a^b p\, G_{xx}(\xi, x) \frac{d^2 u}{dx^2}\, dx = -u(\xi)$$

(4) und

$$\int_a^b p\, G_{\xi xx}(\xi, x) \frac{d^2 u}{dx^2}\, dx = -u'(\xi).$$

Insbesondere ist

$$\int_a^b p\, (G_{xx}(\xi, x))^2\, dx = -G(\xi, \xi) > 0$$

und

$$\int_a^b p (G_{\xi xx}(\xi, x))^2\, dx = -G_{\xi x}(\xi, \xi) > 0.$$

Kehren wir zur Differentialgleichung (1) zurück. Die Theorie der Integralgleichungen liefert vermittelst des symmetrischen Kernes $G(\xi, x)$ die Eigenwerte $\nu_i > 0$ und die Eigenfunktionen $\omega_i(x)$ sowie die Entwickelbar-

keit jeder zulässigen Funktion, deren Ableitungen überdies bis zur vierten Ordnung stetig sind. Für zwei derartige Funktionen $u(x)$ und $v(x)$ gilt

$$\int_a^b p \frac{d^2u}{dx^2} \frac{d^2v}{dx^2} dx = \sum_i \nu_i \int_a^b \omega_i u\, dx \int_a^b \omega_i v\, dx. \tag{5}$$

Es ist nämlich

$$\int_a^b p \frac{d^2u}{dx^2} \frac{d^2v}{dx^2} dx = \int_a^b \frac{d^2}{dx^2}\left(p \frac{d^2u}{dx^2}\right) v\, dx = \sum_i \int_a^b \frac{d^2}{dx^2}\left(p \frac{d^2u}{dx^2}\right) \omega_i dx \int_a^b \omega_i v\, dx$$

$$= \sum_i \int_a^b \frac{d^2}{dx^2}\left(p \frac{d^2\omega_i}{dx^2}\right) u\, dx \int_a^b \omega_i v\, dx.$$

Die Beziehung (5) gilt, wie man zeigen kann, überhaupt für alle Funktionen, für die das Integral linker Hand gebildet werden kann[7]).

Die Klasse der „zulässigen Funktionen" soll alle Funktionen umfassen, die samt ihrer Ableitung stetig sind, der Randbedingung (2) *genügen und für die die Reihe* $\sum_i \nu_i \left(\int_a^b \omega_i\, u\, dx\right)^2$ *konvergiert.* Das Integral $\int_a^b p \left(\frac{d^2u}{dx^2}\right)^2 dx$, für das wir in Anlehnung an eine bekannte Schreibweise zur Abkürzung auch $N\{u\}$ schreiben wollen, können wir uns im Notfalle durch

$$N\{u\} = \int_a^b p \left(\frac{d^2u}{dx^2}\right)^2 dx = \sum_i \nu_i \left(\int_a^b \omega_i u\, dx\right)^2 \tag{5a}$$

definiert denken, und entsprechend das Integral $\int_a^b p \frac{d^2u}{dx^2} \frac{d^2v}{dx^2} dx$ durch (5).

Setzt man in (5) für v die Funktion $G(\xi, x)$ ein, so erhält man unter Berücksichtigung von (4) für alle zulässigen Funktionen u

$$u(\xi) = \sum_i \omega_i(\xi) \int_a^b \omega_i u\, dx, \tag{6}$$

und durch Einsetzen von G_ξ erhält man

$$u'(\xi) = \sum_i \omega_i'(\xi) \int_a^b \omega_i u\, dx. \tag{7}$$

---

[7]) Vgl. z. B. Courant-Hilbert, Methoden der mathematischen Physik, 2. Aufl., S. 43; wäre die Gleichung (5) nicht erfüllt, so würde doch jedenfalls die zugehörige quadratische Relation mit dem Zeichen $\geqq$ statt $=$ gelten, daher steht die Konvergenz der Reihe rechter Hand von vornherein fest.

Die Reihen konvergieren unbedingt und gleichmäßig; denn es gilt vermöge der Schwarzschen Ungleichung

$$\Big(\sum_i \nu_i \int_a^b \omega_i G(\xi,x)\,dx \int_a^b \omega_i u\,dx\Big)^2 \leqq \sum_i \nu_i \Big(\int_a^b \omega_i G(\xi,x)\,dx\Big)^2 \sum_i \nu_i \Big(\int_a^b \omega_i u\,dx\Big)^2$$

$$= N\{u\}\int_a^b p(G_{xx}(\xi,x))^2\,dx = -N\{u\}\,G(\xi,\xi) > 0$$

und die analoge Beziehung für G_ξ. Damit ist zugleich gezeigt: Ist umgekehrt eine Zahlenfolge $x_1, x_2, \ldots$ gegeben und konvergiert die Reihe $\sum_i \nu_i x_i^2$, so konvergieren die Reihen $\sum_i x_i \omega_i(x)$ und $\sum_i x_i \omega_i'(x)$ unbedingt und gleichmäßig und stellen eine zulässige Funktion und ihre Ableitung dar. Jetzt erkennt man, daß wir mit (5a) das Integral $\int_a^b p\left(\frac{d^2u}{dx^2}\right)^2 dx$, wenn man $u = \sum_i x_i \omega_i$ setzt, ganz naturgemäß durch

$$\lim_{n\to\infty} \int_a^b p\Big(\sum_{i=1}^n x_i \frac{d^2\omega_i}{dx^2}\Big)^2 dx$$

definiert haben.

Setzt man für $u(x)$ in (6) die Funktion $G(\xi^*, x)$ und in (7) $G_\xi(\xi^*, x)$ ein, so kommt

$$G(\xi^*,\xi) = -\sum_i \frac{\omega_i(\xi^*)\,\omega_i(\xi)}{\nu_i}$$

und

$$G_{\xi x}(\xi^*,\xi) = -\sum_i \frac{\omega_i'(\xi^*)\,\omega_i'(\xi)}{\nu_i}.$$

Speziell ist

$$G(\xi,\xi) = -\sum_i \frac{\omega_i^2(\xi)}{\nu_i}$$

(8) und

$$G_{\xi x}(\xi,\xi) = -\sum_i \frac{\omega_i'^2(\xi)}{\nu_i},$$

und aus der gleichmäßigen Konvergenz dieser Reihen, welche aus dem Dinischen Satze folgt, schließt man unmittelbar auf die der vorigen in beiden Argumenten.

2. Wir schreiben die allgemeine selbstadjungierte inhomogene lineare Differentialgleichung vierter Ordnung in der Form

$$(9)\qquad \frac{d^2}{dx^2}\Big(p(x)\frac{d^2u}{dx^2}\Big) + \frac{d}{dx}\Big([\lambda^{(1)} r(x) + \lambda^{(2)} r_1(x)]\frac{du}{dx}\Big) - [\lambda^{(3)} k(x) + \lambda^{(4)} k_1(x)]\,u(x) = h(x).$$

$p(x)$ ist die vorhin eingeführte Funktion, $r(x)$ und $r_1(x)$ sind samt ihrer ersten Ableitung in $\langle a\,b\rangle$ stetige Funktionen, $k(x)$, $k_1(x)$ und $h(x)$ in $\langle a\,b\rangle$ stetige Funktionen, $\lambda^{(1)}, \ldots, \lambda^{(4)}$ reelle Parameter, über die später zwecks Aufstellung verschiedener Eigenwertprobleme spezielle Verfügungen getroffen werden sollen. Gesucht wird eine nebst ihren Ableitungen bis zur vierten Ordnung stetige Funktion $u(x)$, die der Differentialgleichung (9) und der Randbedingung (2) genügt.

Wir gehen von folgendem Problem aus, das dem eben formulierten äquivalent ist: für alle zulässigen Funktionen $v(x)$ soll die Gleichung

$$(10)\quad \int_a^b \Big[p\frac{d^2u}{dx^2}\frac{d^2v}{dx^2} - (\lambda^{(1)}r + \lambda^{(2)}r_1)\frac{du}{dx}\frac{dv}{dx} - (\lambda^{(3)}k + \lambda^{(4)}k_1)uv - hv\Big]dx = 0$$

erfüllt sein. Es genügt, wie sich zeigen wird, zu verlangen, daß $u(x)$ eine zulässige Funktion ist. Wir setzen

$$u(x) = \sum_i x_i\omega_i(x), \qquad v(x) = \sum_i y_i\omega_i(x),$$

dann ist

$$u'(x) = \sum_i x_i\omega_i'(x), \qquad v'(x) = \sum_i y_i\omega_i'(x),$$

und sämtliche Reihen konvergieren unbedingt und gleichmäßig. Durch $X_i = \sqrt{\nu_i}\,x_i$, $Y_i = \sqrt{\nu_i}\,y_i$ führen wir Systeme von unendlich vielen Veränderlichen ein, derart, daß zulässige Funktionen und „Vektoren" von beschränktem Betrage einander eindeutig entsprechen. So entstehen aus den Termen der linken Seite von (10), wenn man noch (5) berücksichtigt, die symmetrischen Bilinearformen

$$\int_a^b p\frac{d^2u}{dx^2}\frac{d^2v}{dx^2}dx = \sum_i X_iY_i = (X, Y),$$

$$\int_a^b r\frac{du}{dx}\frac{dv}{dx}dx = \sum_{i,j}\frac{X_iY_j}{\sqrt{\nu_i\nu_j}}\int_a^b r\frac{d\omega_i}{dx}\frac{d\omega_j}{dx}dx = R(X, Y),$$

$$\int_a^b r_1\frac{du}{dx}\frac{dv}{dx}dx = \sum_{i,j}\frac{X_iY_j}{\sqrt{\nu_i\nu_j}}\int_a^b r_1\frac{d\omega_i}{dx}\frac{d\omega_j}{dx}dx = R_1(X, Y),$$

$$\int_a^b kuv\,dx = \sum_{i,j}\frac{X_iY_j}{\sqrt{\nu_i\nu_j}}\int_a^b k\omega_i\omega_j\,dx = K(X, Y),$$

$$\int_a^b k_1uv\,dx = \sum_{i,j}\frac{X_iY_j}{\sqrt{\nu_i\nu_j}}\int_a^b k_1\omega_i\omega_j\,dx = K_1(X, Y)$$

und die Linearform

$$\int_a^b h\, v\, dx = \sum_j \frac{Y_j}{\sqrt{\nu_j}} \int_a^b h\, \omega_j\, dx = H(Y).$$

$H(Y)$ ist beschränkt, und die vier Bilinearformen sind vollstetig[8]). Dies folgt aus der Konvergenz der Summe bzw. der Doppelsummen der Koeffizientenquadrate, welche man wiederum aus der gleichmäßigen Konvergenz der Reihen (8) erschließt. Denn es ist z. B.

$$\left(\int_a^b r\, \omega_i'\, \omega_j'\, dx\right)^2 \leqq \mathrm{Max}\, r^2 \int_a^b \omega_i'^2\, dx \int_a^b \omega_j'^2\, dx,$$

woraus für die Doppelsumme der Koeffizientenquadrate des n-ten Abschnitts von $R(X, Y)$

$$\sum_{i,j=1}^{m} \frac{1}{\nu_i \nu_j} \left(\int_a^b r\, \omega_i'\, \omega_j'\, dx\right)^2 \leqq \mathrm{Max}\, r^2 \left(\int_a^b \sum_{i=1}^{m} \frac{\omega_i'^2}{\nu_i}\, dx\right)^2$$

folgt.

Mit den eingeführten Bezeichnungen geht die Gleichung (10) in

$$(11)\quad (X, Y) - \lambda^{(1)} R(X, Y) - \lambda^{(2)} R_1(X, Y) - \lambda^{(3)} K(X, Y) - \lambda^{(4)} K_1(X, Y) = H(Y)$$

über. Für das hieraus entspringende System von unendlich vielen linearen Gleichungen gilt wegen der Vollstetigkeit der Bilinearformen der Hilbertsche Alternativsatz. Entweder gibt es eine eindeutig bestimmte Lösung von (11), wie man auch die Funktion $h(x)$ gewählt hat, oder die zugehörige homogene Gleichung besitzt Lösungen. In jedem Falle ist $\sum_i X_i^2$ konvergent, und

$$u(x) = \sum_i \frac{X_i}{\sqrt{\nu_i}}\, \omega_i(x)$$

ist eine zulässige Funktion, die entweder die Gleichung (10) erfüllt oder die Gleichung, die aus (10) entsteht, wenn man $h \equiv 0$ setzt. Daß $u(x)$ überdies stetige Ableitungen bis zur vierten Ordnung besitzt und die Differentialgleichung (9) erfüllt (bzw., im anderen Falle, die zugehörige homogene Differentialgleichung), sieht man so ein. Die Beziehung (11) ist insbesondere für alle Wertsysteme (Y_i) erfüllt, die solchen Funktionen $v(x)$ entsprechen, deren Ableitungen bis zur vierten Ordnung stetig sind und bis zur dritten Ordnung sämtlich in a und b verschwinden. Für solche (Y_i) gilt, wie man leicht sieht, jedenfalls

$$(X, Y) = \int_a^b \frac{d^2}{dx^2}\left(p \frac{d^2 v}{dx^2}\right) u\, dx.$$

---

[8]) D. Hilbert, Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen, Leipzig-Berlin 1924, S. 147.

Wir führen die Differentiation aus und formen die entstehenden Terme, sowie alle übrigen Terme der linken Seite von (10) durch teilweise Integration so um, daß (10) die Form

$$\int_a^b N(x)\frac{d^4v}{dx^4}dx = 0$$

erhält. Jetzt folgt in bekannter Weise[9]), daß die so aus $p, r, \ldots, u$ und $\frac{du}{dx}$ gebildete Funktion $N(x)$ eine ganze rationale Funktion dritten Grades ist. Durch sukzessives Differenzieren ergibt sich unsere Behauptung unmittelbar.

3. Zum einfachsten Eigenwertproblem gelangen wir von unserer allgemeinen Differentialgleichung (9) durch die Festsetzung $\lambda^{(1)} = \lambda^{(2)} = \lambda^{(4)} = 0$. Wir betrachten also die Gleichung

$$\frac{d^2}{dx^2}\left(p(x)\frac{d^2u}{dx^2}\right) - \lambda k(x)\,u(x) = 0, \tag{12}$$

bzw.

$$(X, Y) - \lambda K(X, Y) = 0.$$

Die Form $K(X, Y)$ ist abgeschlossen, wenn $k(x)$ in keinem in $\langle a\,b\rangle$ gelegenen Intervall identisch verschwindet[10]). Wir lassen jedoch zunächst die Möglichkeit zu, daß $K(X, Y)$ *nicht* abgeschlossen ist.

Es seien λ_α $(\alpha = 1, 2, \ldots)$ die Eigenwerte, $L^{(\alpha)}(X) = \sum_i l_i^{(\alpha)} X_i$ die normierten Eigenformen von $K(X, Y)$. Man sieht unmittelbar, daß λ_α auch Eigenwert der Differentialgleichung (12) und

$$\psi_\alpha(x) = \sum_i \frac{l_i^{(\alpha)}}{\sqrt{\nu_i}}\omega_i(x) \tag{13}$$

die zugehörige Eigenfunktion ist. Nach dem Früheren ist

$$\psi_\alpha'(x) = \sum_i \frac{l_i^{(\alpha)}}{\sqrt{\nu_i}}\omega_i'(x),$$

und die Reihen rechter Hand konvergieren unbedingt und gleichmäßig.

Bekanntlich ist das Auftreten zweifacher Eigenwerte bei dem vorliegenden Problem nicht ausgeschlossen; aber man sieht leicht, daß die Eigenformen und Eigenfunktionen einander jedenfalls umkehrbar eindeutig entsprechen[11]).

---

[9]) Vgl. z. B. O. Bolza, Vorlesungen über Variationsrechnung, Leipzig-Berlin 1909, S. 153.

[10]) Vgl. L. Lichtenstein, loc. cit.[1]), S. 124.

[11]) Vgl. L. Lichtenstein, Zur Analysis der unendlich vielen Variablen, Zweite Abhandlung, Math. Zeitschr. **3** (1919), S. 139.

Aus den Orthogonalitätsrelationen der Eigenformen folgen für die Eigenfunktionen die Beziehungen

$$\int_a^b p \frac{d^2\psi_\alpha}{dx^2} \frac{d^2\psi_\beta}{dx^2} dx = \delta_{\alpha\beta},$$

welche gleichbedeutend sind mit

$$\int_a^b k \psi_\alpha \psi_\beta dx = \frac{1}{\lambda_\alpha} \delta_{\alpha\beta}.$$

Wir setzen

$$\psi_\alpha(x) = \sqrt{\frac{1}{|\lambda_\alpha|}} \varphi_\alpha(x) \tag{14}$$

und erhalten für die normierten Eigenfunktionen $\varphi_\alpha(x)$ die Orthogonalitätsrelationen

$$\begin{aligned} &\int_a^b k(x)\, \varphi_\alpha(x)\, \varphi_\beta(x)\, dx = 0 \qquad (\alpha \neq \beta), \\ &\int_a^b k(x)\, \varphi_\alpha^2(x)\, dx = \frac{\lambda_\alpha}{|\lambda_\alpha|}. \end{aligned} \tag{15}$$

Für alle Wertsysteme (X_i), (Y_i) mit konvergenter Quadratsumme gilt

$$K(X, Y) = \sum_\alpha \frac{1}{\lambda_\alpha} L^{(\alpha)}(X)\, L^{(\alpha)}(Y).$$

Setzt man

$$\sum_i \frac{X_i}{\sqrt{\nu_i}} \omega_i(x) = u(x), \quad \sum_i \frac{Y_i}{\sqrt{\nu_i}} \omega_i(x) = v(x),$$

so folgt hieraus nach leichter Umformung für alle zulässigen Funktionen $u(x)$ und $v(x)$

$$\int_a^b k\, u\, v\, dx = \sum_\alpha \frac{\lambda_\alpha}{|\lambda_\alpha|} \int_a^b k \varphi_\alpha u\, dx \int_a^b k \varphi_\alpha v\, dx.$$

Setzt man speziell $v(x) = u(x)$, so kommt

$$\int_a^b k u^2 dx = \sum_i \frac{\lambda_\alpha}{|\lambda_\alpha|} \left(\int_a^b k \varphi_\alpha u\, dx \right)^2. \tag{16}$$

Hieraus schließt man auf die Existenz unendlich vieler positiver Eigenwerte, falls $k(x)$ in einem in $\langle a\, b \rangle$ gelegenen Intervall positiv ist. Nimmt $k(x)$ negative Werte an, so gibt es unendlich viele negative Eigenwerte. Endlich gibt es unendlich viele Eigenwerte beiderlei Vorzeichens, falls $k(x)$ das Zeichen wechselt[12]).

[12]) Vgl. loc. cit. [1]), S. 126.

Ist die Form $K(X, Y)$ abgeschlossen, so gilt

$$(X, Y) = \sum_{\alpha} L^{(\alpha)}(X)\, L^{(\alpha)}(Y).$$

Daraus ergibt sich wie vorhin nach leichter Umformung die für alle zulässigen Funktionen $u(x)$ und $v(x)$ gültige Beziehung

$$\int_a^b p \frac{d^2u}{dx^2} \frac{d^2v}{dx^2}\, dx = \sum_{\alpha} |\lambda_\alpha| \int_a^b k\varphi_\alpha u\, dx \int_a^b k\varphi_\alpha u\, dx, \tag{17}$$

und für $v(x) = u(x)$

$$N\{u\} = \int_a^b p \left(\frac{d^2u}{dx^2}\right)^2 dx = \sum_{\alpha} |\lambda_\alpha| \left(\int_a^b k\varphi_\alpha u\, dx\right)^2. \tag{18}$$

Falls $K(X, Y)$ nicht abgeschlossen ist, tritt an die Stelle von (18) die Ungleichheitsbeziehung

$$N\{u\} \geqq \sum_{\alpha} |\lambda_\alpha| \left(\int_a^b k\varphi_\alpha u\, dx\right)^2.$$

Es sei $K(X, Y)$ abgeschlossen, und es möge $f(x)$ eine zulässige Funktion bezeichnen; $f(x)$ besitzt also stetige Ableitung, genügt der Randbedingung (2) und ist in eine Reihe

$$f(x) = \sum_i f_i \omega_i(x)$$

entwickelbar, derart, daß $\sum_i \nu_i f_i^2$ konvergiert, während über die zweite Ableitung keine weitere Voraussetzung gemacht wird, als daß $N\{f\}$ im durch (5a) definierten Sinne existiert. Wir führen in (17) für $u(x)$ die eingangs eingeführte Greensche Funktion $G(\xi, x)$, für $v(x)$ die Funktion $f(x)$ ein. Wegen (4) entsteht aus (17)

$$f(\xi) = \sum_{\alpha} \frac{\lambda_\alpha}{|\lambda_\alpha|} \varphi_\alpha(\xi) \int_a^b k\varphi_\alpha f\, dx, \tag{19}$$

und wenn man anstatt $G(\xi, x)$ die Funktion $G_\xi(\xi, x)$ einführt,

$$f'(\xi) = \sum_{\alpha} \frac{\lambda_\alpha}{|\lambda_\alpha|} \varphi'_\alpha(\xi) \int_a^b k\varphi_\alpha f\, dx. \tag{20}$$

Beide Reihenentwicklungen konvergieren unbedingt und gleichmäßig, wie man genau wie im ersten Abschnitt beweist. Speziell gilt für die Greensche Funktion (man setze $f(x) = G(\xi^*, x)$)

$$G(\xi^*, \xi) = -\sum_{\alpha} \frac{\varphi_\alpha(\xi^*)\, \varphi_\alpha(\xi)}{|\lambda_\alpha|}$$

und

$$G_{\xi x}(\xi^*, \xi) = -\sum_{\alpha} \frac{\varphi'_\alpha(\xi^*)\, \varphi'_\alpha(\xi)}{|\lambda_\alpha|}.$$

4. Für das Folgende und vor allem für die Variationsrechnung ist das spezielle Eigenwertproblem von Nutzen, das sich ergibt, wenn wir $\lambda^{(2)} = \lambda^{(4)} = 0$, $\lambda^{(1)} = \lambda^{(3)} = \lambda$ setzen. Die Differentialgleichung lautet jetzt

$$\frac{d^2}{dx^2}\left(p(x)\frac{d^2u}{dx^2}\right) + \lambda\left[\frac{d}{dx}\left(r(x)\frac{du}{dx}\right) - k(x)\,u(x)\right] = 0, \tag{21}$$

und das lineare Gleichungssystem, wenn wir zur Abkürzung

$$R(X,Y) + K(X,Y) = \mathfrak{K}(X,Y)$$

setzen,

$$(X,Y) - \lambda\,\mathfrak{K}(X,Y) = 0.$$

Die Form $\mathfrak{K}(X,Y)$ ist *nicht* abgeschlossen, wenn es eine zulässige Funktion $u(x)$ gibt, so daß für alle zulässigen Funktionen $v(x)$

$$\int_a^b \left[r\frac{du}{dx}\frac{dv}{dx} + k\,u\,v\right]dx = 0$$

ist. Ist in keinem in $\langle a\,b\rangle$ gelegenen Intervalle $r \equiv 0$ und zugleich $k \equiv 0$, so heißt das: es gibt eine Lösung der Gleichung

$$\frac{d}{dx}\left(r\frac{du}{dx}\right) - k\,u = 0,$$

welche in $\langle a\,b\rangle$ zwei doppelte Nullstellen besitzt. (Außerhalb dieser Nullstellen setzen wir, falls diese nicht die Punkte a und b sind, $u \equiv 0$.) Solches doppelte Verschwinden ist aber nur in Punkten möglich, wo $r = 0$ ist. Man erhält also, falls die hinreichende Bedingung $r = k \equiv 0$ in keinem Intervalle erfüllt ist, die folgende notwendige Bedingung dafür, daß $\mathfrak{K}(X,Y)$ *nicht* abgeschlossen ist: die Funktion $r(x)$ muß in $\langle a\,b\rangle$ mindestens zwei Nullstellen besitzen. Ein Intervall, in dem $r \equiv 0$ ist, zählt dabei als *eine* Nullstelle. Keine der beiden Nullstellen darf im Innern eines Intervalles liegen, in dem $k \equiv 0$ ist, und falls eine auf dem Rande eines solchen Intervalles liegt, so muß dieses Intervall außerhalb des von den beiden Nullstellen gebildeten liegen. Ist diese Bedingung nicht erfüllt, so ist $\mathfrak{K}(X,Y)$ gewiß abgeschlossen. Die folgenden Entwicklungen gelten in jedem Falle.

Wir wollen wieder mit λ_α die Eigenwerte, mit $L^{(\alpha)}(X) = \sum_i l_i^{(\alpha)} X_i$ die Eigenformen diesmal von $\mathfrak{K}(X,Y)$, mit $\varphi_\alpha(x)$ die zugehörigen gemäß (13) und (14) gebildeten normierten Eigenfunktionen bezeichnen. Dann zeigt die Rechnung, daß jetzt folgende Orthogonalitätsrelationen bestehen. Es ist

$$\int_a^b \left[r(x)\frac{d\varphi_\alpha}{dx}\frac{d\varphi_\beta}{dx} + k(x)\,\varphi_\alpha(x)\,\varphi_\beta(x)\right]dx = 0 \qquad (\alpha \neq \beta),$$

$$\int_a^b \left[r(x)\left(\frac{d\varphi_\alpha}{dx}\right)^2 + k(x)\,\varphi_\alpha^2(x)\right]dx = \frac{\lambda_\alpha}{|\lambda_\alpha|}.$$

Der Beziehung

$$\mathfrak{R}(X,Y)=\sum_{\alpha}{}'\frac{1}{\lambda_\alpha}L^{(\alpha)}(X)\,L^{(\alpha)}(Y)$$

entspricht die für alle zulässigen Funktionen gültige Gleichung

$$\int_a^b\Big[r\frac{du}{dx}\frac{dv}{dx}+kuv\Big]dx=\sum_{\alpha}{}'\frac{\lambda_\alpha}{|\lambda_\alpha|}\int_a^b\Big[r\frac{d\varphi_\alpha}{dx}\frac{du}{dx}+k\varphi_\alpha u\Big]dx\int_a^b\Big[r\frac{d\varphi_\alpha}{dx}\frac{dv}{dx}+k\varphi_\alpha v\Big]dx,$$

woraus für *eine* zulässige Funktion

$$\text{(22)}\qquad\int_a^b\Big[r\Big(\frac{du}{dx}\Big)^2+ku^2\Big]dx=\sum_{\alpha}{}'\frac{\lambda_\alpha}{|\lambda_\alpha|}\Big(\int_a^b\Big[r\frac{d\varphi_\alpha}{dx}\frac{du}{dx}+k\varphi_\alpha u\Big]dx\Big)^2$$

folgt. Aus dieser Beziehung folgt ähnlich wie früher, daß es unendlich viele positive, negative oder positive und negative Eigenwerte gibt, je nachdem $r(x)$ und $k(x)$ in Intervallen, die innerhalb $\langle a\,b\rangle$ liegen, beide zugleich positive, negative oder Werte beider Vorzeichen annehmen. Der Fall, wo $r(x)$ und $k(x)$ überall entgegengesetztes Vorzeichen haben, bleibt unentschieden.

Unter der Voraussetzung, daß $\mathfrak{R}(X,Y)$ abgeschlossen ist, folgt aus der dann geltenden Beziehung

$$(X,Y)=\sum_{\alpha}{}'L^{(\alpha)}(X)\,L^{(\alpha)}(Y),$$

daß für alle zulässigen Funktionen $u(x)$ und $v(x)$

$$\int_a^b p\frac{d^2u}{dx^2}\frac{d^2v}{dx^2}dx=\sum_{\alpha}{}'|\lambda_\alpha|\int_a^b\Big[r\frac{d\varphi_\alpha}{dx}\frac{du}{dx}+k\varphi_\alpha u\Big]dx\int_a^b\Big[r\frac{d\varphi_\alpha}{dx}\frac{dv}{dx}+k\varphi_\alpha v\Big]dx$$

ist. In jedem Falle — auch wenn $\mathfrak{R}(X,Y)$ nicht abgeschlossen ist — gilt

$$\text{(24)}\qquad N\{u\}\geqq\sum_{\alpha}{}'|\lambda_\alpha|\Big(\int_a^b\Big[r\frac{d\varphi_\alpha}{dx}\frac{du}{dx}+k\varphi_\alpha u\Big]dx\Big)^2.$$

Es sei $f(x)$ wie im vorigen Abschnitt eine zulässige Funktion. Falls (X,Y) abgeschlossen ist, kann $f(x)$ in eine nach den Eigenfunktionen $\varphi_\alpha(x)$ fortschreitende Reihe entwickelt werden. Man erhält, indem man $\Gamma(\xi,x)$ und $f(x)$ in (23) einsetzt,

$$\text{(25)}\qquad f(\xi)=\sum_{\alpha}{}'\frac{\lambda_\alpha}{|\lambda_\alpha|}\varphi_\alpha(\xi)\int_a^b\Big[r\frac{d\varphi_\alpha}{dx}\frac{df}{dx}+k\varphi_\alpha f\Big]dx.$$

Es ist ferner

$$\text{(26)}\qquad f'(\xi)=\sum_{\alpha}{}'\frac{\lambda_\alpha}{|\lambda_\alpha|}\varphi'_\alpha(\xi)\int_a^b\Big[r\frac{d\varphi_\alpha}{dx}\frac{df}{dx}+k\varphi_\alpha f\Big]dx,$$

und beide Reihen konvergieren unbedingt und gleichmäßig. Die Greensche Funktion $G(\xi^*, \xi)$ gestattet auch nach den Eigenfunktionen dieses Abschnitts eine unbedingt und in beiden Argumenten gleichmäßig konvergente Entwicklung

$$G(\xi^*, \xi) = -\sum_{\alpha} \frac{\varphi_\alpha(\xi^*)\,\varphi_\alpha(\xi)}{|\lambda_\alpha|}.$$

5. Wir setzen jetzt $\lambda^{(1)} = \lambda^{(3)} = \lambda$, $\lambda^{(2)} = \lambda^{(4)} = 1$, behandeln also das Eigenwertproblem der Differentialgleichung

$$(27)\quad \frac{d^2}{dx^2}\left(p(x)\frac{d^2u}{dx^2}\right) + \frac{d}{dx}\left([\lambda r(x) + r_1(x)]\frac{du}{dx}\right) - [\lambda k(x) + k_1(x)]\,u(x) = 0$$

bzw. der Gleichung

$$(X, Y) - \Re(X, Y) - \lambda \mathfrak{R}(X, Y) = 0,$$

wobei zur Abkürzung $R_1(X, Y) + K_1(X, Y) = \Re(X, Y)$ und wie früher $K(X, Y) + R(X, Y) = \mathfrak{R}(X, Y)$ gesetzt wurde.

Es seien μ_α $(\alpha = 1, 2, \ldots)$ die Eigenwerte, $N^{(\alpha)}(X) = \sum_i n_i^{(\alpha)} X_i$ die normierten Eigenformen der quadratischen Form $\Re(X, X)$, und es möge kein Wert μ_α im Intervall $\langle 01\rangle$ liegen. Dies ist z. B. sicher der Fall, wenn wir $r_1(x) \leqq 0$ und $k_1(x) \leqq 0$ voraussetzen. Denn dann sind, wie aus den Entwicklungen des vorigen Abschnitts hervorgeht (vgl. (22)), alle μ_α negativ.

$\Re(X, Y)$ ist im allgemeinen nicht abgeschlossen. Wir ergänzen $N^{(\alpha)}(X)$ durch die Linearformen $M^{(\alpha)}(X)$ zu einem vollständigen Orthogonalsystem. Wegen

$$\Re(X, Y) = \sum_{\alpha} \frac{1}{\mu_\alpha} N^{(\alpha)}(X)\, N^{(\alpha)}(Y)$$

ist

$$X_\alpha^* = \sqrt{1 - \frac{1}{\mu_\alpha}}\, N^{(\alpha)}(X), \qquad {}^*X_\alpha = M^{(\alpha)}(X)$$

eine lineare Transformation, welche

$$(28)\qquad (X, Y) - \Re(X, Y) = (\overline{X}, \overline{Y})$$

liefert, wenn mit $\overline{X}_\alpha$ die irgendwie einfach geordneten X_α^*, ${}^*X_\alpha$ bezeichnet werden; $X_\alpha = \overline{N}^{(\alpha)}(\overline{X})$ sei die zu ihr inverse Transformation. Wir setzen $\overline{\mathfrak{R}}(\overline{X}, \overline{Y}) = \mathfrak{R}(X, Y)$; dann ist $\overline{\mathfrak{R}}$ vollstetig und auch abgeschlossen, wenn $\mathfrak{R}$ es ist. Es seien λ_α die Eigenwerte, $\overline{L}^{(\alpha)}(\overline{X}) = \sum_i \overline{l}_i^{(\alpha)} \overline{X}_i$ die normierten Eigenformen von $\overline{\mathfrak{R}}(\overline{X}, \overline{Y})$. Man sieht sofort, daß

$$\psi_\alpha(x) = \sum_i \frac{l_i^{(\alpha)}}{\sqrt{\nu_i}}\, \omega_i(x),$$

wobei $\overline{N}^{(i)}(\overline{l}^{(\alpha)}) = l_i^{(\alpha)}$ gesetzt ist, Eigenfunktion von (27) zum Eigenwert λ_α ist. Wegen der Beziehung

$$\sum_i \overline{l}_i^{(\alpha)} \overline{l}_i^{(\beta)} = \sum_i l_i^{(\alpha)} l_i^{(\beta)} - \Re(l^{(\alpha)}, l^{(\beta)}),$$

die aus (28) folgt, entsprechen den für die $\overline{L}^{(\alpha)}(X)$ geltenden Orthogonalitätsrelationen die Relationen

$$\int_a^b \left[p \frac{d^2\psi_\alpha}{dx^2} \frac{d^2\psi_\beta}{dx^2} - r_1 \frac{d\psi_\alpha}{dx} \frac{d\psi_\beta}{dx} - k_1 \psi_\alpha \psi_\beta \right] dx = \delta_{\alpha\beta},$$

aus denen nach kurzer Rechnung für die in derselben Weise wie früher (gemäß (14)) eingeführten normierten Eigenfunktionen $\varphi_\alpha(x)$ die Orthogonalitätsrelationen

$$\int_a^b \left[r(x) \frac{d\varphi_\alpha}{dx} \frac{d\varphi_\beta}{dx} + k(x) \varphi_\alpha(x) \varphi_\beta(x) \right] dx = 0 \qquad (\alpha \neq \beta),$$

$$\int_a^b \left[r(x) \left(\frac{d\varphi_\alpha}{dx} \right)^2 + k(x) \varphi_\alpha^2(x) \right] dx = \frac{\lambda_\alpha}{|\lambda_\alpha|}$$

folgen.

Es gilt

$$\overline{\Re}(\overline{X}, \overline{Y}) = \sum_\alpha \frac{1}{\lambda_\alpha} \overline{L}^{(\alpha)}(\overline{X}) \overline{L}^{(\alpha)}(\overline{Y})$$

$$= \sum_\alpha \frac{1}{\lambda_\alpha} [L^{(\alpha)}(X) - \Re(l^{(\alpha)}, X)] [L^{(\alpha)}(Y) - \Re(l^{(\alpha)}, Y)],$$

woraus mit

$$u(x) = \sum_i \frac{X_i}{\sqrt{\nu_i}} \omega_i(x), \qquad v(x) = \sum_i \frac{Y_i}{\sqrt{\nu_i}} \omega_i(x)$$

nach einiger Umformung für alle zulässigen Funktionen

$$\int_a^b \left[r \frac{du}{dx} \frac{dv}{dx} + k u v \right] dx = \sum_\alpha \frac{\lambda_\alpha}{|\lambda_\alpha|} \int_a^b \left[r \frac{d\varphi_\alpha}{dx} \frac{du}{dx} + k \varphi_\alpha u \right] dx \int_a^b \left[r \frac{d\varphi_\alpha}{dx} \frac{dv}{dx} + k \varphi_\alpha v \right] dx$$

folgt. Aus der hieraus durch Identifizierung von $u(x)$ und $v(x)$ sich ergebenden Relation erhält man für das Vorhandensein unendlich vieler positiver, negativer oder positiver und negativer Eigenwerte dieselben Kriterien wie im vorigen Abschnitt.

Jetzt sei vorausgesetzt, daß $\Re(X, Y)$ und damit $\overline{\Re}(\overline{X}, \overline{Y})$ abgeschlossen ist. Es gilt

$$(\overline{X}, \overline{Y}) = \sum_\alpha \overline{L}^{(\alpha)}(\overline{X}) \overline{L}^{(\alpha)}(\overline{Y}),$$

was wegen (28) mit

$$(X,Y)-\Re(X,Y)=\sum_{\alpha}[L^{(\alpha)}(X)-\Re(l^{(\alpha)},X)]\,[L^{(\alpha)}(Y)-\Re(l^{(\alpha)},Y)]$$

gleichbedeutend ist. Daraus folgt für beliebige zulässige Funktionen $u(x)$ und $v(x)$ die Beziehung

$$(29)\quad \int_a^b\Big[p\frac{d^2u}{dx^2}\frac{d^2v}{dx^2}-r_1\frac{du}{dx}\frac{dv}{dx}-k_1uv\Big]dx$$
$$=\sum_{\alpha}|\lambda_\alpha|\int_a^b\Big[r\frac{d\varphi_\alpha}{dx}\frac{du}{dx}+k\varphi_\alpha u\Big]dx\int_a^b\Big[r\frac{d\varphi_\alpha}{dx}\frac{dv}{dx}+k\varphi_\alpha v\Big]dx$$

und die hieraus durch Gleichsetzen von $u(x)$ und $v(x)$ entstehende Gleichung, die im allgemeinen Falle, wenn das Nichtabgeschlossensein von $\Re(X,Y)$ zugelassen wird, so lautet:

$$\int_a^b\Big[p\Big(\frac{d^2u}{dx^2}\Big)^2-r_1\Big(\frac{du}{dx}\Big)^2-k_1u^2\Big]dx\geqq\sum_{\alpha}|\lambda_\alpha|\Big(\int_a^b\Big[r\frac{d\varphi_\alpha}{dx}\frac{du}{dx}+k\varphi_\alpha u\Big]dx\Big)^2.$$

Es sei $\Gamma(\xi,x)$ die zur Differentialgleichung

$$(30)\qquad \frac{d^2}{dx^2}\Big(p\frac{d^2u}{dx^2}\Big)+\frac{d}{dx}\Big(r_1\frac{du}{dx}\Big)-k_1u=0$$

gehörige Greensche Funktion. Sie kann in derselben Weise wie die Funktion $G(\xi,x)$ im ersten Abschnitt gebildet werden; denn wegen unserer Voraussetzung, daß von den Eigenwerten μ_α der kleinste positive >1 ist, existiert keine zulässige Funktion, die die Gleichung (30) erfüllt. Aus den Gleichungen (22) und (24) des vierten Abschnitts können wir nämlich die folgende für alle zulässigen Funktionen geltende Beziehung herleiten:

$$\int_a^b\Big[p\Big(\frac{d^2u}{dx^2}\Big)^2-r_1\Big(\frac{du}{dx}\Big)^2-k_1u^2\Big]dx\geqq\sum_{\alpha}|\mu_\alpha|\Big(1-\frac{1}{\mu_\alpha}\Big)\Big(\int_a^b\Big[r_1\frac{d\chi_\alpha}{dx}\frac{du}{dx}+k_1\chi_\alpha u\Big]dx\Big)^2;$$

hier sind $\chi_\alpha(x)$ die Eigenfunktionen des aus (30) nach Art von (21) gebildeten Eigenwertproblems (dessen Eigenwerte die μ_α sind). Der Ausdruck linker Hand ist demnach für alle zulässigen Funktionen positiv und von Null verschieden. Andererseits folgt aus unseren Entwicklungen die Gleichung

$$\int_a^b\Big[p\Big(\frac{d^2u}{dx^2}\Big)^2-r_1\Big(\frac{du}{dx}\Big)^2-k_1u^2\Big]dx=0,$$

wenn $u(x)$ eine zu unserer Randbedingung für ein in $\langle a\,b\rangle$ enthaltenes Intervall gehörige Lösung von (30) ist, die man außerhalb des Intervalls durch $u\equiv 0$ fortsetzt.

Daher ist das im ersten Abschnitt angegebene Verfahren für die Herstellung von $\Gamma(\xi, x)$ anwendbar; $\Gamma(\xi, x)$ genügt denselben Stetigkeits- und Sprungbedingungen wie $G(\xi, x)$. An die Stelle von (4) tritt die Relation

$$\int_a^b \Big[p\Gamma_{xx}(\xi, x)\frac{d^2u}{dx^2} - r_1\Gamma_x(\xi, x)\frac{du}{dx} - k_1\Gamma(\xi, x)\,u\Big]dx = -u(\xi)$$

und die entsprechende Relation für $\Gamma_\xi(\xi, x)$.

Wir setzen $\Gamma(\xi, x)$ bzw. $\Gamma_\xi(\xi, x)$ und eine willkürliche Funktion $f(x)$, die wie in den vorigen Abschnitten „zulässig" sein muß, in (29) ein und erhalten

$$f(\xi) = \sum_\alpha \frac{\lambda_\alpha}{|\lambda_\alpha|}\varphi_\alpha(\xi)\int_a^b \Big[r\frac{d\varphi_\alpha}{dx}\frac{df}{dx} + k\varphi_\alpha f\Big]dx \tag{31}$$

und

$$f'(\xi) = \sum_\alpha \frac{\lambda_\alpha}{|\lambda_\alpha|}\varphi'_\alpha(\xi)\int_a^b \Big[r\frac{d\varphi_\alpha}{dx}\frac{df}{dx} + k\varphi_\alpha f\Big]dx. \tag{32}$$

Beide Reihen konvergieren unbedingt und gleichmäßig. Wir setzen speziell $f(x) = \Gamma(\xi^*, x)$, dann $f(x) = \Gamma_\xi(\xi^*, x)$ und finden, daß

$$\Gamma(\xi^*, \xi) = -\sum_\alpha \frac{\varphi_\alpha(\xi^*)\varphi_\alpha(\xi)}{|\lambda_\alpha|}$$

und

$$\Gamma_{\xi x}(\xi^*, \xi) = -\sum_\alpha \frac{\varphi'_\alpha(\xi^*)\varphi'_\alpha(\xi)}{|\lambda_\alpha|}$$

ist.

6. Es bleibt noch zu untersuchen, unter welchen Bedingungen eine nach den Eigenfunktionen fortschreitende Reihenentwicklung stattfinden kann, wenn die zugrunde gelegte Bilinearform nicht abgeschlossen ist. Wir betrachten zuerst den Fall des dritten Abschnitts. Es sei also $k \equiv 0$ in gewissen Intervallen innerhalb $\langle a\,b\rangle$. Wenn die Gleichungen (19) und (20) richtig sind, dann sieht man leicht, daß $f(x)$ in den fraglichen Intervallen der Differentialgleichung $\frac{d^2}{dx^2}\left(p\frac{d^2u}{dx^2}\right) = 0$ genügt. Ist andererseits $f(x)$ eine zulässige Funktion, die in den Intervallen, wo $k \equiv 0$ ist, stetige Ableitungen bis zur vierten Ordnung besitzt und diese Gleichung erfüllt, so kann man unschwer zeigen, daß die Entwicklung (19), (20) richtig ist[13]).

Ist in der im fünften Abschnitt behandelten Differentialgleichung (27) $r \equiv 0$, so daß nur der Term, der k enthält, vom Eigenwertparameter abhängt, so liegen ähnliche Verhältnisse vor. Es muß dann die Funktion $f(x)$ in den Intervallen, wo $k \equiv 0$ ist, der Gleichung (30) genügen.

[13]) Vgl. L. Lichtenstein, loc. cit. [1]), S. 134.

Es liege das Problem des vierten Abschnitts vor, und die Form $\mathfrak{R}(X, Y)$ sei nicht abgeschlossen. Wir ergänzen das System der $L^{(\alpha)}(X)$ zu einem vollständigen durch die Linearformen $R^{(\alpha)}(X) = \sum_i r_i^{(\alpha)} X_i$.

Es ist

$$(33) \qquad (X, Y) = \sum_\alpha L^{(\alpha)}(X) L^{(\alpha)}(Y) + \sum_\alpha R^{(\alpha)}(X) R^{(\alpha)}(Y).$$

Es sei $f(x)$ eine zulässige Funktion mit der Entwicklung

$$f(x) = \sum_i f_i \omega_i(x).$$

Bekanntlich konvergiert die Reihe $\sum \nu_i f_i^2$. Aus (33) folgt, daß die Entwicklung (25), (26) dann und nur dann gültig ist, wenn für alle $R^{(\alpha)}(X)$

$$\sum_i \sqrt{\nu_i}\, r_i^{(\alpha)} f_i = 0$$

ist. Analoges gilt im allgemeinen Falle des fünften Abschnitts.

7. Wir wollen die übrigen festen Randbedingungen, die man bei der linearen Differentialgleichung vierter Ordnung stellen kann, außer Betracht lassen und vielmehr noch zwei Randwertaufgaben spezieller Art behandeln, die für die Variationsrechnung interessant sind, und zwar solche, bei denen der Eigenwertparameter in der Randbedingung vorkommt. Wir stellen zunächst die Randbedingung

$$(34) \qquad \begin{gathered} u'(a) = u'(b) = 0, \\ \left[\frac{d}{dx}\left(p \frac{d^2u}{dx^2}\right) - \lambda h_1 u\right]_{x=a} = 0, \qquad \left[\frac{d}{dx}\left(p \frac{d^2u}{dx^2}\right) + \lambda h_2 u\right]_{x=b} = 0 \end{gathered}$$

und schreiben die Differentialgleichung sogleich in der Form, die wir für die Variationsrechnung brauchen werden:

$$(35) \qquad \frac{d^2}{dx^2}\left(p \frac{d^2u}{dx^2}\right) - q u + \lambda\left[\frac{d}{dx}\left(r \frac{du}{dx}\right) - k u\right] = 0.$$

Hier ist $q(x)$ eine beliebige in $\langle a\, b\rangle$ stetige negative Funktion. Die übrigen Zeichen haben dieselbe Bedeutung wie früher.

Es seien $\bar{\nu}_i$ die Eigenwerte, $\bar{\omega}_i(x)$ die normierten Eigenfunktionen der Differentialgleichung

$$(36) \qquad \frac{d^2}{dx^2}\left(p \frac{d^2u}{dx^2}\right) - q u - \nu u = 0,$$

die zur Randbedingung

$$(37) \quad u'(a) = u'(b) = 0, \qquad \frac{d}{dx}\left(p \frac{d^2u}{dx^2}\right) = 0 \quad \text{für } x = a \text{ und } x = b$$

gehören. Da $q(x) < 0$ ist, existiert die zu dieser Randbedingung gehörende Greensche Funktion $\bar{G}(\xi, x)$ der Differentialgleichung

$$\frac{d^2}{dx^2}\left(p \frac{d^2u}{dx^2}\right) - q u = 0$$

und kann wie im ersten Abschnitt zur Herleitung der Eigenwerte und Eigenfunktionen von (36) und zum Beweise eines zu dem dortigen analogen Entwicklungssatzes verwandt werden. „Zulässig" heißt jetzt jede stetige Funktion $u(x)$ mit stetiger erster und quadratisch integrierbarer zweiter Ableitung, die der Bedingung $u'(a) = u'(b) = 0$ genügt. Für zulässige Funktionen gilt

$$\int_a^b \left[p\, \bar{G}_{xx}(\xi, x)\, \frac{d^2 u}{dx^2} - q\, \bar{G}(\xi, x)\, u \right] dx = -u(\xi),$$

$$\int_a^b \left[p\, \bar{G}_{\xi xx}(\xi, x)\, \frac{d^2 u}{dx^2} - q\, \bar{G}_{\xi}(\xi, x)\, u \right] dx = -u'(\xi).$$

Ferner erhält man für Funktionen mit stetigen Ableitungen bis zur vierten Ordnung, die der Randbedingung (37) genügen, und danach für alle zulässigen Funktionen die Relation

$$\int_a^b \left[p\, \frac{d^2 u}{dx^2}\, \frac{d^2 v}{dx^2} - q\, u\, v \right] dx = \sum_i \int_a^b \left[p\, \frac{d^2 \bar{\omega}_i}{dx^2}\, \frac{d^2 u}{dx^2} - q\, \bar{\omega}_i u \right] dx \int_a^b \bar{\omega}_i\, v\, dx$$

$$= \sum_i \bar{\nu}_i \int_a^b \bar{\omega}_i\, u\, dx \int_a^b \bar{\omega}_i\, v\, dx$$

und durch Einsetzen der Greenschen Funktion die differenzierbare Reihenentwicklung für alle zulässigen Funktionen.

Wir kehren zur Differentialgleichung (35) mit der Randbedingung (34) zurück und ersetzen sie durch die äquivalente Gleichung

$$\int_a^b \left[p\, \frac{d^2 u}{dx^2}\, \frac{d^2 v}{dx^2} - q\, u\, v \right] dx$$

$$- \lambda \left\{ \int_a^b \left[r\, \frac{du}{dx}\, \frac{dv}{dx} + k\, u\, v \right] dx + h_1\, u(a)\, v(a) + h_2\, u(b)\, v(b) \right\} = 0,$$

die für alle zulässigen Funktionen $v(x)$ erfüllt sein soll. Wir setzen

$$u(x) = \sum_i \frac{X_i}{\sqrt{\bar{\nu}_i}}\, \bar{\omega}_i(x), \qquad v(x) = \sum_i \frac{Y_i}{\sqrt{\bar{\nu}_i}}\, \bar{\omega}_i(x)$$

und erhalten

$$(X, Y) - \lambda K(X, Y) = 0$$

mit der vollstetigen symmetrischen Bilinearform

$$K(X, Y) = \sum_{i,j} \frac{X_i Y_i}{\sqrt{\bar{\nu}_i \bar{\nu}_j}} \left\{ \int_a^b \left[r\, \frac{d\bar{\omega}_i}{dx}\, \frac{d\bar{\omega}_j}{dx} + k\, \bar{\omega}_i \bar{\omega}_j \right] dx + h_1\, \bar{\omega}_i(a)\, \bar{\omega}_j(a) + h_2\, \bar{\omega}_i(b)\, \bar{\omega}_j(b) \right\}.$$

(Die Konvergenz z. B. der Doppelreihe $\sum_{i,j} \frac{\bar{\omega}_i^2(a)\bar{\omega}_j^2(a)}{\bar{\nu}_i \bar{\nu}_j}$ folgt aus der gleichmäßigen Konvergenz von $\sum_i \frac{\bar{\omega}_i^2}{\bar{\nu}_i} = -\bar{G}(x,x)$).

Es seien λ_α die Eigenwerte, $L^{(\alpha)}(X) = \sum_i l_i^{(\alpha)} X_i$ die Eigenformen von $K(X, Y)$; dann sind λ_α die Eigenwerte und

$$\varphi_\alpha(x) = \sqrt{|\lambda_\alpha|} \sum_i \frac{l_i^{(\alpha)}}{\sqrt{\bar{\nu}_i}} \bar{\omega}_i(x)$$

die normierten Eigenfunktionen von (35), (34), für die folgende Orthogonalitätsbeziehungen gelten:

$$\int_a^b \left[r \frac{d\varphi_\alpha}{dx} \frac{d\varphi_\beta}{dx} + k \varphi_\alpha \varphi_\beta \right] dx + h_1 \varphi_\alpha(a) \varphi_\beta(a) + h_2 \varphi_\alpha(b) \varphi_\beta(b) = 0 \quad (\alpha \neq \beta),$$

$$\int_a^b \left[r \left(\frac{d\varphi_\alpha}{dx} \right)^2 + k \varphi_\alpha^2 \right] dx + h_1 \varphi_\alpha^2(a) + h_2 \varphi_\alpha^2(b) = \frac{\lambda_\alpha}{|\lambda_\alpha|}.$$

Aus

$$K(X, X) = \sum_\alpha \frac{1}{\lambda_\alpha} [L^{(\alpha)}(X)]^2$$

folgt für alle zulässigen Funktionen

$$(38) \qquad \int_a^b \left[r \left(\frac{du}{dx} \right)^2 + k u^2 \right] dx + h_1 u^2(a) + h_2 u^2(b)$$

$$= \sum_\alpha \frac{\lambda_\alpha}{|\lambda_\alpha|} \left\{ \int_a^b \left[r \frac{d\varphi_\alpha}{dx} \frac{du}{dx} + k \varphi_\alpha u \right] dx + h_1 \varphi_\alpha(a) u(a) + h_2 \varphi_\alpha(b) u(b) \right\}^2.$$

Die Form $K(X, Y)$ ist abgeschlossen, wenn es keine zulässige Funktion $u(x)$ gibt, so daß für alle zulässigen Funktionen

$$\int_a^b \left[r \frac{du}{dx} \frac{dv}{dx} + k u v \right] dx + h_1 u(a) v(a) + h_2 u(b) v(b) = 0$$

ist. Daraus ergeben sich, wie man leicht sieht, die gleichen Kriterien wie im vierten Abschnitt; denn jene Funktion $u(x)$ muß gewiß am Rande verschwinden.

Aus

$$(X, X) \geqq \sum_\alpha [L^{(\alpha)}(X)]^2$$

folgt für alle zulässigen Funktionen

$$(39)\qquad \int_a^b \left[p\left(\frac{d^2u}{dx^2}\right)^2 - q\,u^2\right]dx$$

$$\geqq \sum_\alpha |\lambda_\alpha| \left\{\int_a^b \left[r\frac{d\varphi_\alpha}{dx}\frac{du}{dx} + k\,\varphi_\alpha u\right]dx + h_1\varphi_\alpha(a)\,u(a) + h_2\varphi_\alpha(b)\,u(b)\right\}^2.$$

Wenn $K(X, Y)$ abgeschlossen ist, so gilt hier das Gleichheitszeichen, und in diesem Falle erhält man durch Einsetzen von $\bar{G}(\xi, x)$ bzw. $\bar{G}_\xi(\xi, x)$ in die zugehörige bilineare Relation den Satz, daß sich jede zulässige Funktion $f(x)$ in eine unbedingt und gleichmäßig konvergente und gliedweise differenzierbare Reihe nach den Eigenfunktionen $\varphi_\alpha(x)$ entwickeln läßt, deren Koeffizienten

$$(40)\quad \frac{\lambda_\alpha}{|\lambda_\alpha|}x_\alpha = \frac{\lambda_\alpha}{|\lambda_\alpha|}\left\{\int_a^b \left[r\frac{d\varphi_\alpha}{dx}\frac{df}{dx} + k\,\varphi_\alpha f\right]dx + h_1\varphi_\alpha(a)\,f(a) + h_2\varphi_\alpha(b)\,f(b)\right\}$$

lauten.

8. Zu ähnlichen Ergebnissen wie die vorige führt die folgende Randwertaufgabe:

$$u(a) = u(b) = 0,$$

$$\left[p\frac{d^2u}{dx^2} + \lambda h_1\frac{du}{dx}\right]_{x=a} = 0, \qquad \left[p\frac{d^2u}{dx^2} - \lambda h_2\frac{du}{dx}\right]_{x=b} = 0.$$

Hierbei können wir die Differentialgleichung ganz in der Form des vierten Abschnitts schreiben:

$$(21)\qquad \frac{d^2}{dx^2}\left(p\frac{d^2u}{dx^2}\right) + \lambda\left[\frac{d}{dx}\left(r\frac{du}{dx}\right) - k\,u\right] = 0.$$

Wir legen jetzt Größen $\overline{\overline{G}}(\xi, x)$, $\bar{\bar{\nu}}_i$ und $\bar{\bar{\omega}}_i(x)$ zugrunde, die wörtlich denselben Relationen genügen wie im ersten Abschnitt die Größen G, ω_i und ν_i. Nur lautet jetzt die Randbedingung: $\bar{\bar{\omega}}_i$ und $\frac{d^2\bar{\bar{\omega}}_i}{dx^2}$ soll am Rande verschwinden, und eine stetige Funktion $u(x)$ heißt „zulässig", wenn sie in a und b verschwindet und stetige Ableitung erster und quadratisch integrierbare Ableitung zweiter Ordnung besitzt. Die zweite Ableitung braucht also der Randbedingung nicht zu genügen. Statt (21) schreiben wir

$$\int_a^b p\frac{d^2u}{dx^2}\frac{d^2v}{dx^2}dx - \lambda\left\{\int_a^b \left[r\frac{du}{dx}\frac{dv}{dx} + k\,u\,v\right]dx + h_1u'(a)v'(a) + h_2u'(b)v'(b)\right\} = 0$$

und erhalten mit

$$u(x) = \sum_i \frac{X_i}{\sqrt{\bar{\bar{\nu}}_i}}\bar{\bar{\omega}}_i(x), \qquad v(x) = \sum_i \frac{Y_i}{\sqrt{\bar{\bar{\nu}}_i}}\bar{\bar{\omega}}_i(x)$$

die Gleichung

$$(X, Y) - \lambda K(X, Y) = 0,$$

wobei

$$K(X,Y)=\sum_{i,j}\frac{X_i Y_j}{\sqrt{\bar{\nu}_i\bar{\nu}_j}}\left\{\int_a^b\left[r\frac{d\bar{\omega}_i}{dx}\frac{d\bar{\omega}_j}{dx}+k\bar{\omega}_i\bar{\omega}_j\right]dx+h_1\bar{\omega}_i'(a)\bar{\omega}_j'(a)+h_2\bar{\omega}_i'(b)\bar{\omega}_j'(b)\right\}$$

wiederum symmetrisch und vollstetig ist. Die Vollstetigkeit folgt in ähnlicher Weise wie im vorigen Abschnitt aus der Konvergenz der Reihe $\sum_i \frac{\bar{\omega}_i'^2}{\sqrt{\bar{\nu}_i}}$.

Die Ableitung der Eigenwerte λ_α und Eigenfunktionen $\varphi_\alpha(x)$ erfolgt wie im vorigen Abschnitt. Es gelten die Orthogonalitätsrelationen

$$\int_a^b\left[r\frac{d\varphi_\alpha}{dx}\frac{d\varphi_\beta}{dx}+k\varphi_\alpha\varphi_\beta\right]dx+h_1\varphi_\alpha'(a)\varphi_\beta'(a)+h_2\varphi_\alpha'(b)\varphi_\beta'(b)=0\quad(\alpha\neq\beta),$$

$$\int_a^b\left[r\left(\frac{d\varphi_\alpha}{dx}\right)^2+k\varphi_\alpha^2\right]dx+h_1\varphi_\alpha'^2(a)+h_2\varphi_\alpha'^2(b)=\frac{\lambda_\alpha}{|\lambda_\alpha|}.$$

Wenn $u(x)$ eine zulässige Funktion ist, so gilt

$$\int_a^b\left[r\left(\frac{du}{dx}\right)^2+ku^2\right]dx+h_1u'^2(a)+h_2u'^2(b) \tag{41}$$

$$=\sum_\alpha\frac{\lambda_\alpha}{|\lambda_\alpha|}\left\{\int_a^b\left[r\frac{d\varphi_\alpha}{dx}\frac{du}{dx}+k\varphi_\alpha u\right]dx+h_1\varphi_\alpha'(a)u'(a)+h_2\varphi_\alpha'(b)u'(b)\right\}^2,$$

und es ist

$$\int_a^b p\left(\frac{d^2u}{dx^2}\right)^2dx \tag{42}$$

$$\geqq\sum_\alpha|\lambda_\alpha|\left\{\int_a^b\left[r\frac{d\varphi_\alpha}{dx}\frac{du}{dx}+k\varphi_\alpha u\right]dx+h_1\varphi_\alpha'(a)u'(a)+h_2\varphi_\alpha'(b)u'(b)\right\}^2,$$

wobei, falls $K(X,Y)$ abgeschlossen ist, stets das Gleichheitszeichen gilt. In diesem Falle läßt sich jede zulässige Funktion $f(x)$ in eine gleichmäßig konvergente und gliedweise differenzierbare Reihe nach den $\varphi_\alpha(x)$ mit den Koeffizienten

$$\frac{\lambda_\alpha}{|\lambda_\alpha|}x_\alpha=\frac{\lambda_\alpha}{|\lambda_\alpha|}\left\{\int_a^b\left[r\frac{d\varphi_\alpha}{dx}\frac{df}{dx}+k\varphi_\alpha f\right]dx+h_1\varphi_\alpha'(a)f'(a)+h_2\varphi_\alpha'(b)f'(b)\right\}^2 \tag{43}$$

entwickeln.

9. Die in den Abschnitten 1 bis 6 benutzten Methoden und Beweisführungen lassen sich unmittelbar auf den Fall der allgemeinen selbstadjungierten linearen Differentialgleichung von der Ordnung $2n$ übertragen.

Wir haben uns bei den ausführlichen Darstellungen der übersichtlicheren Schreibweise halber auf den Fall $n = 2$ beschränkt. Hier seien noch die wesentlichsten Formeln und Ergebnisse angegeben, die man im allgemeinen Falle erhält.

Die Randbedingung sei

$$(44)\quad u(a) = u(b) = u'(a) = u'(b) = \ldots = u^{(n-1)}(a) = u^{(n-1)}(b) = 0;$$

„zulässige" Funktionen haben stetige Ableitungen bis zur Ordnung $n-1$, quadratisch integrierbare (analog (5a)) Ableitung n-ter Ordnung und genügen der Randbedingung. $p(x)$ sei eine stetige positive Funktion mit stetigen Ableitungen bis zur Ordnung n. Durch die zur Randbedingung (44) gehörige Greensche Funktion $G(\xi, x)$ der Differentialgleichung

$$\frac{d^n}{dx^n}\left(p(x)\frac{d^n u}{dx^n}\right) = 0,$$

die, samt ihren Ableitungen [14]) bis zur Ordnung $2n-2$ stetig, der Sprungbedingung

$$\left[\frac{\partial^{n-1}}{\partial x^{n-1}}\left(p\frac{\partial^n G(\xi, x)}{\partial x^n}\right)\right]_{x=\xi-0}^{x=\xi+0} = -1$$

genügt und so nach der Methode des ersten Abschnitts gewonnen werden kann, und für die die Relationen

$$\int_a^b p\frac{\partial^n G(\xi, x)}{\partial x^n}\frac{d^n u}{dx^n}\,dx = -(-1)^n u(\xi),$$

$$\int_a^b p\frac{\partial^n}{\partial x^n}\left(\frac{\partial^\nu G(\xi, x)}{\partial \xi^\nu}\right)\frac{d^n u}{dx^n}\,dx = -(-1)^n u^{(\nu)}(\xi) \qquad (\nu = 1, \ldots, n-1)$$

gelten, erhält man die (positiven) Eigenwerte ν_i und die Eigenfunktionen $\omega_i(x)$ der Differentialgleichung

$$\frac{d^n}{dx^n}\left(p\frac{d^n u}{dx^n}\right) - (-1)^n \nu u = 0$$

und die für sie gültigen Entwicklungssätze. Die Formel (5) lautet jetzt

$$\int_a^b p\frac{d^n u}{dx^n}\frac{d^n v}{dx^n}\,dx = \sum_i \nu_i \int \omega_i u\,dx \int \omega_i v\,dx,$$

und indem man hierin für $v(x)$ die Greensche Funktion und ihre Ableitungen nach ξ einsetzt, erhält man die gleichmäßig konvergente Reihenentwicklung für eine zulässige Funktion $u(x)$, welche $n-1$ mal gliedweise

[14]) Unter „Ableitung von der Ordnung $n+\sigma$" die Größe $\frac{\partial^\sigma}{\partial x^\sigma}\left(p\frac{\partial^n G}{\partial x^n}\right)$ verstanden.

differenziert werden darf. Setzt man dann $u = G(\xi, x)$, $\frac{\partial G(\xi, x)}{\partial \xi}, \ldots, \frac{\partial^{n-1} G}{\partial \xi^{n-1}}$, so erhält man die für uns wichtige gleichmäßige Konvergenz der Reihen

$$\sum_i \frac{\omega_i^2}{\nu_i}, \quad \sum_i \frac{\omega_i'^2}{\nu_i}, \quad \ldots, \quad \sum_i \frac{\omega_i^{(n-1)^2}}{\nu_i}. \tag{45}$$

Wir betrachten die Differentialgleichung

$$\sum_{\sigma=0}^{n} (-1)^\sigma \frac{d^{n-\sigma}}{dx^{n-\sigma}} \left(p_\sigma(x) \frac{d^{n-\sigma} u}{dx^{n-\sigma}} \right) = 0. \tag{46}$$

Die Funktionen $p_\sigma(x)$ haben stetige Ableitungen bis zur Ordnung $n - \sigma$. p_0 ist positiv. Um zu dem für die Variationsrechnung brauchbaren Eigenwertproblem zu kommen, setzen wir

$$p_0 = p, \quad p_\sigma = -\lambda r_\sigma \ (\sigma = 1, 2, \ldots, n-1), \quad p_n = -\lambda k$$

und gehen von der Beziehung

$$\int_a^b \left[p \frac{d^n u}{dx^n} \frac{d^n v}{dx^n} - \lambda r_1 \frac{d^{n-1} u}{dx^{n-1}} \frac{d^{n-1} v}{dx^{n-1}} - \ldots - \lambda r_{n-1} \frac{du}{dx} \frac{dv}{dx} - \lambda k u v \right] dx = 0$$

aus, in der wir

$$u = \sum_i \frac{X_i}{\sqrt{\nu_i}} \omega_i, \qquad v = \sum_i \frac{Y_i}{\sqrt{\nu_i}} \omega_i$$

setzen. Die Vollstetigkeit der Bilinearformen

$$\sum_{i,j} \frac{X_i Y_j}{\sqrt{\nu_i \nu_j}} \int_a^b r_\sigma \frac{d^{n-\sigma} \omega_i}{dx^{n-\sigma}} \frac{d^{n-\sigma} \omega_j}{dx^{n-\sigma}} dx$$

folgt wegen

$$\sum_{i,j=1}^{m} \frac{1}{\nu_i \nu_j} \left(\int_a^b r_\sigma \frac{d^{n-\sigma} \omega_i}{dx^{n-\sigma}} \frac{d^{n-\sigma} \omega_j}{dx^{n-\sigma}} dx \right)^2 \leqq \operatorname{Max} r_\sigma^2 \sum_{i,j=1}^{m} \frac{1}{\nu_i \nu_j} \int_a^b \left(\frac{d^{n-\sigma} \omega_i}{dx^{n-\sigma}} \right)^2 dx \int_a^b \left(\frac{d^{n-\sigma} \omega_j}{dx^{n-\sigma}} \right)^2 dx$$

$$= \operatorname{Max} r_\sigma^2 \left(\int_a^b \sum_{i=1}^{m} \frac{\omega_i^{(n-\sigma)^2}}{\nu_i} dx \right)^2$$

aus der gleichmäßigen Konvergenz der Reihen (45). Wir erhalten wieder eine Gleichung

$$(X, Y) - \lambda \mathfrak{R}(X, Y) = 0,$$

aus der alles Weitere folgt. Die Eigenwerte λ_α der Form $\mathfrak{R}(X, Y)$ sind zugleich die der Differentialgleichung, und aus den Eigenformen erhält man in der üblichen Weise deren normierte Eigenfunktionen $\varphi_\alpha(x)$. Die

Orthogonalitätsbeziehungen lauten

$$\int_a^b \left[r_1 \frac{d^{n-1}\varphi_\alpha}{dx^{n-1}} \frac{d^{n-1}\varphi_\beta}{dx^{n-1}} + \ldots + r_{n-1} \frac{d\varphi_\alpha}{dx} \frac{d\varphi_\beta}{dx} + k\varphi_\alpha\varphi_\beta \right] dx = \frac{\lambda_\alpha}{|\lambda_\alpha|} \delta_{\alpha\beta}.$$

Für jede zulässige Funktion $u(x)$ bestehen, mit

$$x_\alpha = \int_a^b \left[r_1 \frac{d^{n-1}\varphi_\alpha}{dx^{n-1}} \frac{d^{n-1}u}{dx^{n-1}} + \ldots + r_{n-1} \frac{d\varphi_\alpha}{dx} \frac{du}{dx} + k\varphi_\alpha u \right] dx,$$

die Beziehungen

$$\int_a^b \left[r_1 \left(\frac{d^{n-1}u}{dx^{n-1}}\right)^2 + \ldots + r_{n-1} \left(\frac{du}{dx}\right)^2 + ku^2 \right] dx = \sum_\alpha \frac{\lambda_\alpha}{|\lambda_\alpha|} x_\alpha^2 \tag{47}$$

und

$$\int_a^b p \left(\frac{d^n u}{dx^n}\right)^2 dx \geqq \sum_\alpha |\lambda_\alpha| x_\alpha^2 \tag{48}$$

und, falls $\mathfrak{R}(X, Y)$ abgeschlossen ist, die Reihenentwicklung

$$u(x) = \sum_\alpha \frac{\lambda_\alpha}{|\lambda_\alpha|} x_\alpha \varphi_\alpha(x).$$

Die Reihe kann $n-1$ mal gliedweise differenziert werden, und alle so erhaltenen Reihen konvergieren absolut und gleichmäßig, was man wie immer durch Einführen der Funktionen $G(\xi, x)$, $\frac{\partial G}{\partial \xi}, \ldots, \frac{\partial^{n-1} G}{\partial \xi^{n-1}}$ in die zur Gleichung (48), in der dann das Gleichheitszeichen gilt, gehörige bilineare Relation bestätigt.

Man könnte noch nach dem Vorbild des fünften Abschnitts das Eigenwertproblem betrachten, das aus (46) entsteht, wenn man die p_σ in der Art $\lambda r_\sigma + s_\sigma$ vom Eigenwertparameter abhängen läßt. Es müssen dann über die Funktionen s_σ gewisse einschränkende Voraussetzungen gemacht werden.

10. Es sei $N(x, t)$ eine im Quadrat $a \leqq x \leqq b$, $a \leqq t \leqq b$ erklärte stetige Funktion, die der Bedingung $N(x, t) = N(t, x)$ genügt; die übrigen Zeichen haben die Bedeutung der Abschnitte 1 und 2. Wir betrachten die Integro-Differentialgleichung

$$\frac{d^2}{dx^2}\left(p \frac{d^2 u}{dx^2}\right) + \lambda \frac{d}{dx}\left(r \frac{du}{dx}\right) - \lambda k u - \lambda \int_a^b N(x, t)\, u(t)\, dt = 0 \tag{49}$$

und die Randbedingung (2). Wir suchen eine zulässige Funktion $u(x)$ (im Sinne des ersten Abschnitts), die für alle zulässigen Funktionen $v(x)$

der Gleichung

$$\int_a^b \left[p \frac{d^2u}{dx^2} \frac{d^2v}{dx^2} - \lambda r \frac{du}{dx} \frac{dv}{dx} - \lambda k u v \right] dx - \lambda \int_a^b \int_a^b N(x,t)\, u(x)\, v(t)\, dx\, dt = 0$$

genügt. Diese Funktion besitzt dann stetige Ableitungen bis zur vierten Ordnung und genügt der Integro-Differentialgleichung (49), wie man in ähnlicher Weise wie früher beweist.

Mit

$$u(x) = \sum \frac{X_i}{\sqrt{\nu_i}} \omega_i(x), \qquad v(x) = \sum \frac{Y_i}{\sqrt{\nu_i}} \omega_i(x)$$

(alle Bezeichnungen sind die der Abschnitte 1 und 2) erhält man sofort

$$(X, Y) - \lambda S(X, Y) = 0,$$

wobei

$$R(X, Y) + K(X, Y) + N(X, Y) = S(X, Y)$$

gesetzt ist und

$$N(X, Y) = \sum_{i,j} \frac{X_i Y_j}{\sqrt{\nu_i \nu_j}} \int_a^b \int_a^b N(x,t)\, \omega_i(x)\, \omega_j(t)\, dx\, dt$$

eine, wie man leicht sieht, vollstetige symmetrische Bilinearform ist.

Aus den Eigenformen und Eigenwerten von $S(X, Y)$ wird in der bekannten Weise das System der Eigenwerte λ_α und normierten Eigenfunktionen $\varphi_\alpha(x)$ von (49) abgeleitet. Im Unterschied zur Sachlage bei der Differentialgleichung lassen sich jetzt über Anzahl, Vorzeichen und Vielfachheit der Eigenwerte nur die Angaben machen, die immer bei vollstetigen symmetrischen Bilinearformen richtig sind: es gibt unendlich viele Eigenwerte, die sich im Endlichen nicht häufen und von denen jeder eine endliche Vielfachheit besitzt. Die Orthogonalitätsbeziehungen lauten

$$\int_a^b \left[r \frac{d\varphi_\alpha}{dx} \frac{d\varphi_\beta}{dx} + k \varphi_\alpha \varphi_\beta \right] dx + \int_a^b \int_a^b N(x,t)\, \varphi_\alpha(x)\, \varphi_\beta(t)\, dx\, dt = \frac{\lambda_\alpha}{|\lambda_\alpha|} \delta_{\alpha\beta},$$

und es gelten die Beziehungen

$$(50) \qquad \int_a^b \left[r \left(\frac{du}{dx} \right)^2 + k u^2 \right] dx + \int_a^b \int_a^b N(x,t)\, u(x)\, u(t)\, dx\, dt = \sum_\alpha \frac{\lambda_\alpha}{|\lambda_\alpha|} x_\alpha^2$$

und

$$(51) \qquad N\{u\} = \int_a^b p \left(\frac{d^2u}{dx^2} \right)^2 dx \geqq \sum_\alpha |\lambda_\alpha| x_\alpha^2,$$

wobei

$$(52) \quad x_\alpha = \int_a^b \left[r \frac{d\varphi_\alpha}{dx} \frac{du}{dx} + k \varphi_\alpha u \right] dx + \int_a^b \int_a^b N(x,t)\, \varphi_\alpha(x)\, u(t)\, dx\, dt$$

gesetzt ist. In (51) gilt immer das Gleichheitszeichen, falls $S(X, Y)$ abgeschlossen ist; $S(X, Y)$ ist *nicht* abgeschlossen, wenn es eine zulässige Funktion $u(x)$ gibt, die für alle zulässigen Funktionen $v(x)$

$$\int_a^b \left[r \frac{du}{dx} \frac{dv}{dx} + k u v \right] dx + \int_a^b \int_a^b N(x,t)\, u(x)\, v(t)\, dx\, dt = 0$$

liefert, mithin der Integro-Differentialgleichung

$$\frac{d}{dx}\left(r \frac{du}{dx} \right) - k u - \int_a^b N(x,t)\, u(t)\, dt$$

genügt. Insbesondere ist $S(X, Y)$ nicht abgeschlossen, wenn es in $\langle a\, b \rangle$ ein Intervall $\langle a_1\, b_1 \rangle$ gibt, so daß $r \equiv 0$ und $k \equiv 0$ in $\langle a_1\, b_1 \rangle$ und $N(x, t) \equiv 0$ in den Streifen $\begin{pmatrix} a_1 \leqq x \leqq b_1 \\ a \leqq t \leqq b \end{pmatrix}$, $\begin{pmatrix} a_1 \leqq t \leqq b_1 \\ a \leqq x \leqq b \end{pmatrix}$ gilt.

Ist $S(X, Y)$ abgeschlossen, so folgt durch Einsetzen der im ersten Abschnitt definierten Greenschen Funktion $G(\xi, x)$ in die zu (51) gehörige bilineare Relation die für jede zulässige Funktion geltende Reihenentwicklung

$$u(x) = \sum_\alpha \frac{\lambda_\alpha}{|\lambda_\alpha|} x_\alpha \varphi_\alpha(x),$$

$$u'(x) = \sum_\alpha \frac{\lambda_\alpha}{|\lambda_\alpha|} x_\alpha \varphi'_\alpha(x),$$

wobei die Koeffizienten x_α durch (52) gegeben sind.

(Eingegangen am 2. März 1931.)

Mitteilungen.

Dem von dem Internationalen Mathematikerkongreß in Bologna 1928 gefaßten Beschluß entsprechend wird vom 4. bis 12. September 1932 in Zürich ein

Internationaler Mathematikerkongreß

stattfinden, auf den das Organisationskomite schon jetzt hinweist. Die ersten genaueren Mitteilungen und Einladungen werden im Oktober 1931 verschickt.

Das vorgesehene wissenschaftliche Programm besteht aus einer größeren Reihe von allgemeinen Vorträgen, durch die ein möglichst vollständiges Bild des gegenwärtigen Standes der Mathematik gegeben werden soll, und aus Sektionssitzungen, die für kürzere Mitteilungen über Ergebnisse neuer Untersuchungen bestimmt sind.

Neben diesen Fachveranstaltungen sind gesellschaftliche Zusammenkünfte und Ausflüge, für die die Schweiz besonders schöne Gelegenheiten bietet, geplant, die, wie wir hoffen, zur Förderung und Vertiefung der wissenschaftlichen und persönlichen Beziehungen zwischen den hier versammelten Mathematikern aller Länder beitragen werden.

Zürich, Sommer 1931. Das Organisationskomitee.

Die Göttinger Gesellschaft der Wissenschaften hat neuerdings einen Vertrieb von Sonderabdrücken eingerichtet, der nach Fachgruppen geordnet ist. Die Fachgruppen der mathematisch-physikalischen Klasse sind folgende: 1. Mathematik, 2. Physik, Astronomie, Technik, 3. Chemie, einschl. Physikalische Chemie, 4. Geologie und Mikroskopie, 5. Geographie und Geophysik, 6. Biologie.

Grundlagen einer Genealogie der periodischen Bahnen im restringierten Dreikörperproblem.

Erste Mitteilung.

Beweis des E. Strömgrenschen dynamischen Abschlußprinzips der periodischen Bahngruppen.

Von

Aurel Wintner in Baltimore, Md.

Inhalt.

Seite

Einleitung 324
§ 1. Vorbereitungen 327
§ 2. Scharparameter und Jacobische Konstante 331
§ 3. E. Strömgrens dynamisches Abschlußprinzip 334
§ 4. Beispiele zum Abschlußprinzip 337
§ 5. Beweis des Abschlußsatzes mittels analytischer Fortsetzung 339
§ 6. Anwendung auf die Thieleschen Bewegungsgleichungen 347

Die numerischen Ergebnisse der Kopenhagener Sternwarte sind, sofern sie bereits veröffentlicht vorliegen (ein Teil des Materials ist z. Z. noch unveröffentlicht), hauptsächlich in den Astr. Nachr. und in den Kopenhagener Akademieberichten (Math.-fys. Meddelelser det Kgl. Danske Videnskabernes Selskab) und außerdem — mit Ausnahme der weiter unten mit III, V, VI bezeichneten Aufsätze — alle auch als besondere Veröffentlichungen der Sternwarte erschienen (Publikationer og mindre Meddelelser fra Københavns Observatorium). Da sie für den Mathematiker wohl vor allem in ihrem vollen Zusammenhang und ohne die numerischen Einzelheiten in Betracht kommen, so sei hier nur auf die bisher erschienenen zusammenfassenden Berichte von E. Strömgren verwiesen, zumal diejenigen leitenden Gesichtspunkte der Kopenhagener numerischen Rechnungen, um welche es sich in der vorliegenden Arbeit handeln wird, nicht in den Veröffentlichungen der Einzeluntersuchungen, sondern nur in diesen [die seither erschienenen Veröffentlichungen, so z. B. die nichtinfinitesimalen Librationen[1]) allerdings nur teilweise berücksichtigenden] Aufsätzen herausgearbeitet sind:

[1]) Vgl. Fußnote [7]) weiter unten.

I. Die auf der Kopenhagener Sternwarte ausgeführten numerischen Untersuchungen über das Dreikörperproblem (= Publ. Observ. Nr. 27) [veröffentlicht von J. Fischer-Petersen], Sonderheft des „Sirius", 1917.

II. Forms of periodic motion in the Restricted Problem and in the general Problem of Three Bodies, according to researches executed at the Observatory of Copenhagen [= Publ. Observ. Nr. 39], Verhandlungen des fünften skandinavischen Mathematiker-Kongresses (1922) in Helsingfors.

III. Tre Aartier Celest Mekanik paa Københavns Observatorium, Saertryk af Universitetets Festskrift, November 1923 (Kapitel V und Bilag).

IV. Zur Durchmusterung des problème restreint (= Publ. Observ. Nr. 47), Seeliger-Festschrift, 1924.

V. Unsere Kenntnisse über die Bewegungsformen im Dreikörperproblem, Ergebnisse der exakten Naturwissenschaften, herausgegeben von der Schriftleitung der „Naturwissenschaften", Bd. 4, 1925.

VI. Das System der periodischen und asymptotischen Bahnen im problème restreint, Vierteljahrsschr. der Astr. Ges. **61** (1926), S. 275–277.

Zur Orientierung sei bemerkt, daß, was numerische Untersuchungen über *periodische Lösungen „im Großen"* mittels numerischer Quadraturen anlangt, außer diesen unter Leitung von E. Strömgren ausgeführten Rechnungen sowie den ersten Ansätzen bei Hill und Thiele (und in gewissem Sinne bei v. Haerdtl) — wenn man etwa von dem leider unzulänglichen zusammenfassenden letzten Abschnitt des Sammelwerkes „Periodic Orbits" (Washington, Carnegie Inst., 1920) von Herrn Moulton absieht — trotz der fast unüberblickbaren Literatur über das Dreikörperproblem eigentlich nur die Rechnungen von Burrau und Darwin zu erwähnen sind, über deren leitende Gesichtspunkte Burrau zusammenfassend berichtet hat [außer in der weiter unten in der Fußnote [16]) zitierten Note] in dem Aufsatz:

VII. Vierteljahresschr. der Astr. Ges. **33** (1898), S. 21–33.

Es handelt sich dabei um die folgenden Arbeiten:

VIII. C. Burrau, Recherches numériques concernant des solutions périodiques d'un cas spécial du problème des trois corps, Astr. Nachr. **135** (1894), S. 233–240; **136** (1894), S. 161–174.

IX. G. H. Darwin. Scientific Papers, Cambridge, **4** (1911), Part. I.

Darwins große Arbeit „Periodic Orbits" (1897) erschien zuerst in den Acta Math. und gleich danach mit einigen Kürzungen in den Math. Annalen; die nachstehenden Zitate beziehen sich auf seine gesammelten Werke (IX), in welchen ein außer von Burrau auch von Poincaré (vgl. S. 352 ff. von XIV weiter unten) und Hough (loc. cit. IX, S. 114 ff.) bemerktes Versehen bereits berücksichtigt ist und welche außerdem einige bereits unter dem Einfluß von Burrau stehende Zusätze enthalten (vgl. insb. loc. cit. IX, S. 141 ff.). In der ersten Veröffentlichung von Darwin fehlt nämlich noch jedes genealogische Prinzip, während Thiele (vgl. loc. cit. VIII, S. 171 oder 234) schon Jahre vorher die fundamentale Idee hatte, daß man eine Ejektionsbahn erreichen könnte und eine diesbezügliche numerische Preisaufgabe stellt, die Burrau gleich löst (VIII). Burrau sagt hier: „A mon sens, voici la véritable règle: les diverses classes de solutions sont séparées les unes des autres par des trajéctoires d'éjection", mit welchem Prinzip dann Darwin jene Versehen nachträglich behebt; in seiner ersten Arbeit hat Darwin an die Erreichung einer Ejektionsbahn noch gar nicht gedacht. —

Das in der vorliegenden Arbeit bewiesene Strömgrensche Prinzip geht freilich über die vorher erwähnte Thielesche Vermutung bzw. über das soeben zitierte Burrausche Prinzip, die bisher mathematisch ebenfalls unbewiesen waren, ganz wesentlich hinaus und schließt diese als notwendige Bestandteile in sich, ja korrigiert sie grundsätzlich dahin, daß es sich nicht um *diverses* classes, sondern um eine und dieselbe Gruppe handelt. Übrigens war jenes Prinzip jahrzehntelang eben nur ein Prinzip, das, mit Ausnahme seiner Anwendung auf die erwähnten Unstimmigkeiten bei Darwin, bei keiner Gruppe realisiert worden war, so daß eine Stoßbahn, bis zur Wiederaufnahme des Problems durch Strömgren, praktisch als ein eventuell provisorischer, eventuell aber endgültiger Gruppenabschluß galt (ja man hat nicht einmal neue provisorische Abschlußbahnen gefunden).

Bisher war aber, wie gesagt, auch dieses engere Prinzip samt der Thieleschen Vermutung mathematisch eine offene Frage, d. h. es bestand für den Mathematiker die Möglichkeit, daß eine Gruppe „ganz ohne Grund" (z. B. wegen einer im *reellen* Gebiete nicht näher interpretierbaren *komplexen* Singularität der Gruppenmannigfaltigkeit) zu existieren aufhört.

In drei kürzlich erschienenen Aufsätzen habe ich begonnen, über die numerischen Ergebnisse eine mathematische Rechenschaft zu geben:

X. Über eine Revision der Sortentheorie des restringierten Dreikörperproblems, Berichte der math.-phys. Klasse der Sächsischen Akademie der Wissenschaften zu Leipzig, **82** (1930), S. 3–56.

XI. Librationstheorie des restringierten Dreikörperproblems (= Publ. Observ. Nr. 71), Math. Zeitschr. **32** (1930), S. 668–674, insb. S. 671.

XII. Über die v. Haerdtlsche Bahn, Math. Zeitschr. **32** (1930), S. 668–674, insb. S. 671.

In X sind u. a. die recht spärlichen Beziehungen des Sortenproblems zum Gruppenproblem bereits besprochen worden und das dort über die Anschmiegungsmethoden Gesagte will ich hier nicht wiederholen; vgl. auch XI, § 4. Was die Problematik der Sorten bzw. der Genren anlangt, sei hier außer

XIII. H. Poincaré, Les méthodes nouvelles de la mécanique céleste, Tome 1 (1892);

XIV. H. Poincaré, Les méthodes nouvelles de la mécanique céleste, Tome 3 (1899);

XV. H. Poincaré, Les solutions périodiques et les planètes du type d'Hécube, Bull. Astr. **19** (1902), S. 177–198,

sowie der in dieser Richtung grundlegenden Untersuchung von

XVI. G. D. Birkhoff, The Restricted Problem of Three Bodies, Rend. Palermo **39** (1915), S. 265–334,

nur die sich an XVI anschließende Arbeit von

XVII. B. O. Koopman, On rejection to infinity and exterior motion in the restricted problem of three bodies, Trans. of Amer. Math. Soc. **29** (1927), S. 287–331

erwähnt [wegen der zweiten Sorte und des zweiten Genres von Poincaré s. insb. XIII, S. 139–144; XIV, S. 201–361, insb. S. 343 ff. und S. 355; ferner XV, S. 181 ff. und

XVIII. K. Schwarzschild, Über eine Classe periodischer Lösungen des Dreikörperproblems, Astr. Nachr. **147** (1896), S. 17–24,

aber auch

XIX. P. Stäckel, Periodische Lösungen zweiter Art beim Dreikörperproblem, Jahresber. des Deutschen Math.-Ver. 28 (1919), S. 180–181

und (auch was XVIII anlangt) X, S. 46–51, S. 29 ff.; ferner XVI, S. 328–333, XVII, S. 323 ff.].

Bemerkung.

Was ich Kopenhagener *Gruppe* nenne, heißt in II „group", in III (dänisch) „Gruppe", in den deutschen Veröffentlichungen der Kopenhagener Sternwarte aber (und auch in V) „Klasse", in VI allerdings abwechselnd „Klasse" und „Gruppe". Ich habe mich für das Wort Gruppe entschieden, damit keine Verwechslung mit den Poincaréschen Begriffen der Sorte sowie des Genres entsteht, die in der deutschen Literatur bei jedem Autor anders heißen (Art, Gattung usw., u. a. — und vor allem — auch Klasse). Die von mir vorgeschlagene Terminologie (sorte = Sorte, genre = Genre, group = Gruppe) ist freilich nicht besonders schön, sie beugt aber Mißverständnissen von vornherein vor. Leider hat bereits Poincaré keine konsequente Terminologie, da er die Darwinsche Labilitätseinteilung in zwei kinds (IX, S. 36 ff.) ebenfalls mit dem Worte sorte wiedergibt (XIV, S. 343–344). Unter Sorte verstehe ich freilich das, was in XIII eindeutig sorte heißt und Darwins kind heißt in X einfach „Art". Endlich nennt Poincaré Darwins kind loc. cit. zwar sorte, in einem anderen Abschnitt (S. 292) ebenfalls von XIV aber classe, so daß eine Verwendung des Wortes Klasse — in einem Zusammenhang, wobei es sich, wie in der vorliegenden Arbeit, nicht nur um Gruppen handelt — so lange irreführend sein dürfte, als sich keine einheitliche Terminologie herausbildet.

Einleitung.

Dieser Aufsatz ist der erste in einer Folge von Arbeiten, die eine analytische Genealogie der periodischen Bahnen des restringierten Dreikörperproblems aufbauen und dabei u. a. eine mathematische Begründung der numerischen Ergebnisse der Kopenhagener Sternwarte erbringen. In dieser vorläufigen Mitteilung wird, unter Ausschaltung aller anderen Probleme (Doppelbahnen, Stabilitätsaustausch, Poincarésche Sorten- bzw. Genreproblematik und sich daran anschließende Birkhoffsche Fragestellungen usf.), für welche durch diese erste Arbeit eine feste Basis gewonnen ist, der im Titel angegebene Satz bewiesen.

Für den Nachweis der Existenz des *Anfangs*stückes einer Gruppe von periodischen Lösungen stehen heute verschiedene Methoden zur Verfügung, durch welche z. B. die Existenz der „dynamisch von vornherein selbstverständlichen" (d. h. ganz in der Nähe der Librationszentra oder der Massen oder des unendlich fernen Punktes verlaufenden) Anfangsbahnen bei allen Kopenhagener Gruppen von a bis m mit Ausnahme von k, und zwar für alle Werte des Massenprozentsatzes, mathematisch bereits sichergestellt ist[2]). Über den *weiteren* (d. h. außerhalb eines irgendwie ermittelten

[2]) Vgl. X und XI; daselbst auch Literatur. Die römischen Ziffern weisen stets auf die auf S. 322 ff. angegebenen Arbeiten hin.

Gruppenstückes erfolgenden) Entwicklungsgang einer Gruppe, ja überhaupt über die Möglichkeit einer solchen Weiterentwicklung war bisher numerisch sehr viel, mathematisch aber gar nichts bekannt [die mathematisch grundlegenden Untersuchungen von Birkhoff und seiner Mitarbeiter beziehen sich entweder[3]) nicht auf das Gruppen-, sondern auf das Sortenproblem (vgl. hierzu S. 349 weiter unten), oder[4]) sie liegen in von diesen beiden verschiedenen Richtungen]. Eine bekannte Stelle der Méthodes Nouvelles[5]) enthält — implizite — allerdings eine Aussage, die sich auf das „Problem im Großen" bezieht, indem sie behauptet, daß periodische Lösungen nur paarweise verschwinden können. Doch läßt sich der Poincarésche Beweis dieses sogenannten Verzweigungssatzes (samt den dabei verwendeten nicht ganz richtig festgelegten Begriffen, wovon aber hier bis S. 335 der Einfachheit halber abgesehen werden soll) nicht aufrechterhalten: daß es sich überhaupt um ein Problem im Großen handelt, wird — und zwar an der Stelle „... au moment où μ passera a une valeur quelconque μ_0" — offenbar übersehen. Es ist nämlich bereits ein Problem im Großen — und deshalb und nur deshalb enthält der Poincarésche Verzweigungssatz eine Aussage im Großen —, *ob* μ einen von Poincaré unbegründeterweise von vornherein fest gewählten Wert μ_0 *überhaupt erreichen*, geschweige denn passieren kann [μ ist dabei irgendein Scharparameter und wäre daher in dem Gruppenproblem, genau umgekehrt wie im Sortenproblem, nach beliebiger, aber fester Wahl des Massenprozentsatzes mit der (relativen) Energie, d. h. mit der Jacobischen Konstante zu identifizieren].

---

[3]) Siehe vor allem XVI, XVII. Vgl. auch H. Hotelling, Trans. of Amer. Math. Soc. 27 (1925), S. 329—344.

[4]) Vgl. vor allem G. D. Birkhoff, Dynamical problems with two degrees of freedom, Trans. of Amer. Math. Soc. 18 (1917), S. 199—300 und M. Morse, The foundations of a theory in the calculus of variations in the large, Trans. of Amer. Math. Soc. 30 (1928), S. 213ff., sowie den daselbst zitierten Beweis der Morseschen Relation.

Für solche Probleme dritter Art ist die gleichzeitige Fixierung der Jacobischen Konstante und des Massenprozentsatzes charakteristisch. Bei den Gruppen und bei den Sorten ist nur je einer dieser beiden Parameter festgehalten, es wäre daher wünschenswert, diese (sich auf *diskrete* Bahnen beziehenden) Ergebnisse dritter Art durch Zugrundelegung der in X benutzten Phasenmannigfaltigkeiten („Massenschranke" usf.) in die *Kontinuum*-Problematik der Gruppen und der Sorten einzuordnen.

Es ist vielleicht nicht überflüssig zu bemerken, daß die Sätze, die Herr Morse in seinen weit ausholenden Untersuchungen als Existenzsätze im Großen bezeichnet, sich entweder auf *eine feste* periodische Lösung oder auf Extremumeigenschaften einer *gegebenen* Schar von periodischen Lösungen beziehen, nicht aber Fortsetzbarkeits- bzw. Existenzaussagen enthalten und daher mit den obigen Strömgrenschen Fragestellungen keine direkten Berührungspunkte zu haben scheinen.

[5]) XIII, S. 83.

Über das Problem im Großen hat nun E. Strömgren[6]) sein über den Poincaréschen Verzweigungssatz noch ganz wesentlich hinausgehendes Prinzip ausgesprochen und — mit Ausnahme der Gruppe g, über welche die Rechnungen z. Z. noch nicht ganz abgeschlossen sind — im Falle gleicher Massen, worauf sich fast alle Kopenhagener Rechnungen beziehen, bei allen seinen Gruppen numerisch bestätigen können. Das Strömgrensche Prinzip besagt, daß jede Gruppe periodischer Lösungen entweder eine in sich geschlossene Mannigfaltigkeit bildet oder ein *natürliches*, dynamisch bedingtes Abklingen erleidet. Dabei wird das Auftauchen einer periodischen Stoßbahn (innerhalb der Gruppe) noch keinesfalls als ein *natürlicher*, endgültiger Gruppenabschluß betrachtet; als ein solcher soll vielmehr nur die unbegrenzte Annäherung an eine Gleichgewichtslösung (d. h. an eine der fünf Lagrangeschen Unbestimmtheitsstellen des regularisierten Phasenraumes) und außerdem das Unendlichwerden entweder der Bahndimensionen oder der Energie (= Jacobische Konstante) oder endlich der Periode gelten [ob dabei die Periode in der gewöhnlichen oder aber in der Thieleschen Zeit gemessen wird, läuft, wie man sich leicht überzeugt, auf dasselbe hinaus]; in dem Strömgrenschen Prinzip nehmen also die im Laufe der Weiterentwicklung eventuell auftauchenden, jedenfalls diskreten Stoßbahnen der Gruppe in dem Kontinuum der Nichtstoßbahnen eine auf keine Weise bevorzugte Stellung ein. — Was das Unendlichwerden der erwähnten drei Bestimmungsstücke betrifft, so ist freilich erlaubt, daß gleichzeitig deren zwei oder auch alle drei unendlich werden, doch gibt es Gruppen, bei welchen nur ein einziges Bestimmungsstück unendlich wird. Dies ist z. B. der Fall, wenn die periodische Gruppe durch eine asymptotische Grenzbahn abgeschlossen wird[7]), da in diesem Falle die Jacobische Konstante und die Bahndimensionen bei Annäherung der Gruppenbahnen an die Grenzbahn beschränkt bleiben können und nur die Periode unendlich wird. Wohlbemerkt, es wird durch das Strömgrensche Prinzip nicht behauptet, daß eine Fortsetzung der Gruppe über ein Unendlichwerden hindurch stets unmöglich sein muß, es dürfen vielmehr z. B. von einer asymptotischen Grenzbahn auch „zwei" Gruppen ausgehen (in diesem Falle wird übrigens auch eine zweite asymptotische Grenzbahn zu erwarten sein); vgl. übrigens § 4. — Der asymptotische Gruppenabschluß kann freilich nicht derart verstanden werden, daß die Gruppenbahnen gegen eine feste asymptotische Kurve *konvergieren*.

In dieser vorläufigen Mitteilung wird nun die Richtigkeit des Strömgrenschen Prinzips für jeden Wert des Massenprozentsatzes ganz allgemein für jede Gruppe bewiesen, die überhaupt existiert, d. h. für welche die Existenz

[6]) IV, V, VI.

[7]) Außer IV vgl. insb. E. Strömgren, Publ. Observ. Nr. 61, 64, 67, 70 (1929—30).

irgendeines beliebig kleinen Gruppenstückes [das nicht notwendig ein „dynamisch von vornherein selbstverständliches" Gruppenstück (vgl. S. 324 oben) zu sein braucht — ein solches würde es z. B. zur Gruppe k gar nicht geben] irgendwie mathematisch bewiesen werden kann, m. a. W. für jede Gruppe, für welche ein „Problem im Kleinen" gelöst ist; es handelt sich dabei um eine analytische und zwar *reell* bleibende Fortsetzung des gegebenen Gruppenstückes, bei welcher, ebenso wie in den numerischen Rechnungen, die Energie der Scharparameter ist, nach welchem bei irgendwie, aber fest gewähltem Massenprozentsatz die diesmal *analytische Fortsetzung* (bei Strömgren etwa „numerische Fortsetzung") der übrigen Integrationskonstanten (also auch der Periode) erfolgt.

* * *

Diese Arbeit sowie die unmittelbar folgende zweite Mitteilung sind im Prinzip im Sommer 1930 während meines Aufenthaltes an der Kopenhagener Sternwarte entstanden. Die bisherigen mathematischen Untersuchungen verfolgten im allgemeinen andere Zwecke, und ohne das numerische Kopenhagener Material wäre man wohl nie zu den umfassenden Strömgrenschen Feststellungen gekommen. So möchte ich Herrn Professor Strömgren auch hier für die vielen Besprechungen, in welchen ich die großzügigen theoretischen Gesichtspunkte, die ihn bei seinen numerischen Untersuchungen leiteten, näher kennenlernen durfte, meinen herzlichsten Dank aussprechen.

§ 1.

Vorbereitungen.

Es seien

$$G_i = G_i(\gamma; \xi_1, \ldots, \xi_n) \qquad (i = 1, \ldots, n) \tag{1}$$

n überall konvergente Potenzreihen, d. h. ganze Funktionen von $n+1$ Veränderlichen. Das Differentialsystem

$$\frac{d\xi_i}{d\tau} = G_i(\gamma; \xi_1, \ldots, \xi_n) \qquad (i = 1, \ldots, n) \tag{2}$$

enthält also einen Parameter γ, so daß eine Lösung von (2) erst durch die $n+1$ Bestimmungsstücke

$$\gamma; \xi_1^0, \ldots, \xi_n^0 \tag{3}$$

festgelegt wird; hierbei ist

$$\xi_j^0 = (\xi_j)_{\tau=0} \qquad (j = 1, \ldots, n) \tag{4}$$

gesetzt. Wir nehmen an, daß es zu (1) eine $(n+1)$-te ganze Funktion

$$\Gamma = \Gamma(\gamma; \xi_1, \ldots, \xi_n) \tag{5}$$

gibt, die die Eigenschaft hat, daß

$$\Gamma(\gamma; \xi_1, \ldots, \xi_n) = 0 \tag{6}$$

eine sogenannte invariante Relation von (2) ist. Dies besagt: Wenn bei gegebenem γ die n Anfangswerte (4) derart gewählt werden, daß sie die Gleichung (6) befriedigen:

$$\Gamma(\gamma; \xi_1^0, \ldots, \xi_n^0) = 0, \tag{7}$$

und setzt man die durch (3) festgelegte Lösung $\xi_i = \xi_i(\tau)$ von (2) in die Funktion (5) ein, so verschwindet die so entstehende Funktion von τ nicht nur für $\tau = 0$, sondern für alle τ. Z. B. ist (6) jedenfalls dann eine invariante Relation, wenn $\Gamma = C$ ($C =$ konst.) ein Integral von (2) ist, und es handelt sich dann darum, daß bei gegebenem γ nur solche Anfangswerte (4) zugelassen werden, für welche die Integrationskonstante C verschwindet (dies ist der Fall in der Regularisierungstheorie des restringierten Dreikörperproblems und allgemeiner[7a]) beim Prinzip von Maupertuis). Ist die ganze Funktion Γ identisch gleich Null, was wir zunächst noch nicht ausschließen wollen, so ist (6) bei jedem (2) eine invariante Relation, so daß wir bisher eigentlich nur eine einzige Annahme gemacht haben, nämlich daß in (2) die Funktionen (1) ganze Funktionen sind. Die durch die Bestimmungsstücke (3) festgelegte Lösung von (2) sei

$$\xi_i = \xi_i(\tau; \gamma; \xi_1^0, \ldots, \xi_n^0) \qquad (i = 1, \ldots, n). \tag{8}$$

Wir nehmen ein für allemal an, daß (8) nicht unabhängig von t ist, m. a. W. daß (4) keine „Unbestimmtheitsstelle" des Phasenraumes darstellt (in dem Falle des restringierten Dreikörperproblems besagt dies, daß (8) nicht eine der fünf Lagrangeschen Gleichgewichtslösungen ist, für die auch das Maupertuissche Prinzip bekanntlich illusorisch wird).

Es sei

$$\hat{\gamma}; \hat{\xi}_1^0, \ldots, \hat{\xi}_n^0 \tag{9}$$

ein fest gewähltes spezielles Wertsystem (3) und

$$\hat{\xi}_i(\tau) = \xi_i(\tau; \hat{\gamma}; \hat{\xi}_1^0, \ldots, \hat{\xi}_n^0) \tag{10}$$

die durch (9) festgelegte Lösung. Es sei bekannt, daß die Funktionen $\hat{\xi}_i(\tau)$ in dem endlichen, abgeschlossenen Intervall

$$0 \leqq \tau \leqq L \qquad (L > 0) \tag{11}$$

regulär und dem Betrage nach

$$\leqq M \qquad (M < +\infty) \tag{12}$$

[7a]) Vgl. G. D. Birkhoff, a. a. O. [4]).

sind, und außerdem sollen (nicht nur die n Größen $\hat{\xi}_j^0$, sondern) alle $n+1$ Größen (9) absolut $\leqq M$ sein. Es sei $\hat{T}$ irgendeine, der Bedingung

$$0 \leqq \hat{T} \leqq L \tag{13}$$

genügende Zahl. Man setze bei beliebiger (also auch komplexer) Wahl von τ, γ und ξ_j^0

$$\vartheta = \tau - \hat{T}, \quad \Delta\gamma = \gamma - \hat{\gamma}, \quad \Delta\xi_j^0 = \xi_j^0 - \hat{\xi}_j^0. \tag{14}$$

Nach bekannten Sätzen gibt es dann zu jedem $\varepsilon > 0$ ein außer von ε nur von L, M und den G_i abhängiges, also von der speziellen Wahl von $\hat{T}, \hat{\gamma}, \hat{\xi}_j^0$ unabhängiges $\delta = \delta_\varepsilon > 0$ derart, daß, sobald die $n+1$ Abweichungen

$$\Delta\gamma = \gamma - \hat{\gamma}, \quad \Delta\xi_j^0 = \xi_j^0 - \hat{\xi}_j^0 \tag{15}$$

in den Kreisbereichen

$$|\Delta\gamma| \leqq \delta_\varepsilon, \quad |\Delta\xi_j^0| \leqq \delta_\varepsilon \tag{16}$$

liegen, die Funktionen (8) in dem Intervalle[8])

$$0 \leqq \tau \leqq L + \varepsilon \tag{17}$$

regulär sind, und daß darüber hinaus die unter Beachtung von (15) durch

$$f_i(\vartheta, \Delta\gamma, \Delta\xi_1^0, \ldots, \Delta\xi_n^0) \equiv \xi_i(\vartheta + \hat{T}; \gamma; \xi_1^0, \ldots, \xi_n^0) \tag{18}$$

definierten Funktionen f_i (die freilich auch von den fest gedachten Parametern $\hat{T}, \hat{\gamma}, \hat{\xi}_j^0$ abhängen) in dem Bereiche

$$|\vartheta| \leqq \varepsilon, \quad |\Delta\gamma| \leqq \delta_\varepsilon, \quad |\Delta\xi_j^0| = \delta_\varepsilon \tag{19}$$

ihrer $n+2$ Argumente (14) regulär und dem Betrage nach z. B.

$$\leqq 2M \tag{20}$$

sind. Setzt man also zunächst $\hat{T} = \vartheta = 0$, so folgt durch Subtraktion bei jedem der Bedingung (13) genügenden $\hat{T}$, daß die unter Beachtung von (15) durch

$$F_i(\vartheta, \Delta\gamma, \Delta\xi_1^0, \ldots, \Delta\xi_n^0) \equiv \xi_i(\vartheta + \hat{T}; \gamma; \xi_1^0, \ldots, \xi_n^0) - \xi_i(0; \gamma; \xi_1^0, \ldots, \xi_n^0) \tag{21}$$

definierten Funktionen F_i (die wieder auch von den fest gedachten Parametern

$$\hat{T}, \quad \hat{\gamma}, \quad \hat{\xi}_1^0, \ldots, \hat{\xi}_n^0 \tag{22}$$

abhängen) in dem Bereiche (19) regulär und dem Betrage nach

$$\leqq 4M \tag{23}$$

sind.

[8]) Man beachte, daß unter (11) auch $\tau = L$ zugelassen worden ist.

Es sei hier notiert, daß wegen (14) und (21) — und mit Rücksicht darauf, daß τ nicht explizit in (2) eintritt — das Erfülltsein der n Bedingungsgleichungen

$$F_i(\vartheta, \Delta\gamma, \Delta\xi_1^0, \ldots, \Delta\xi_n^0) = 0 \qquad (i = 1, \ldots, n) \tag{24}$$

notwendig und hinreichend ist dafür, daß die Lösung (8) eine für $-\infty < \tau < +\infty$ periodische Lösung ist, die die Periode

$$\hat{T} + \vartheta \tag{25}$$

hat (und zwar muß dann ϑ reell sein). — Es handelt sich also nur um die wohlbekannte Poincarésche Methode[9]).

Wir wollen noch, bei einer gegebenen ganzen Funktion (5) und bei einem fest gewählten Wertsystem (22), unter Beachtung von (14) eine Funktion

$$F_0 = F_0(\vartheta, \Delta\gamma, \Delta\xi_1^0, \ldots, \Delta\xi_n^0) \tag{26}$$

einführen, derart, daß

$$F_0(\vartheta, \Delta\gamma, \Delta\xi_1^0, \ldots, \Delta\xi_n^0) = 0 \tag{27}$$

die notwendige und hinreichende Bedingung darstellt dafür, daß durch die Lösung (8) zur Zeit $\tau = \vartheta + \hat{T}$ die Gleichung (6) befriedigt wird. Wir beschränken uns gleich auf den Fall, daß (5) eine invariante Relation von (2) ist. Dann ist (26) unabhängig von ϑ (und auch von $\hat{T}$), und (27) ist die notwendige und hinreichende Bedingung für (7). Die Funktion (26) wird daher einfach dadurch erhalten, daß man die Potenzreihe

$$\Gamma(\gamma; \xi_1^0, \ldots, \xi_n^0) \equiv \Gamma(\hat{\gamma} + \Delta\gamma; \hat{\xi}_1^0 + \Delta\xi_1^0, \ldots, \hat{\xi}_n^0 + \Delta\xi_n^0) \tag{28}$$

bei festem $\hat{\gamma}$ und $\hat{\xi}_j^0$ nach Potenzen von $\Delta\gamma$ und der $\Delta\xi_j^0$ ordnet. (26) ist also eine ganze Funktion ihrer angeschriebenen Argumente, und was deren erstes, nämlich ϑ anlangt, so gilt sogar

$$\frac{\partial F_0}{\partial \vartheta} \equiv 0. \tag{29}$$

Da es für § 5 (S. 339 w. u.) sehr wesentlich ist, so sei hier wiederholt (vgl. S. 328 oben), daß das „Maupertuissche Integral“ (6) mit dem eventuellen Integrale $\Gamma = \text{konst.} \gtreqless 0$ von (2) nicht zu verwechseln ist, wenn γ in (2), (6) tatsächlich eingeht; denn (7) wird dann nicht von selbst, sondern nur dann erfüllt sein, wenn man in $\Gamma = \text{konst.}$ die rechter Hand stehende Integrationskonstante gleich Null wählt. Auch sei hier die oben (S. 328) eingeführte Voraussetzung wiederholt, wonach (8) und daher auch (10) keine Gleichgewichtslösung ist.

---

[9]) XIII, S. 79ff.

§ 2.

Scharparameter und Jacobische Konstante.

Wir setzen von jetzt an voraus, daß die $n+1$ Funktionen (1), (5) für reelle Werte ihrer Argumente reellwertig und daß auch die $n+1$ Zahlen (9) reell sind, und wollen nur Wertsysteme (9) in Betracht ziehen, die der Bedingung $\Gamma = 0$ genügen:

$$\Gamma(\hat{\gamma}^0, \hat{\xi}_1^0, \ldots, \hat{\xi}_n^0) = 0, \tag{30}$$

so daß also die invariante Relation (6) durch die zu (9) gehörige Lösung (10) von (2) für alle τ befriedigt wird. Und zwar soll (9) derart sein, daß (10) für reelle τ eine Periode hat, die wir mit $\hat{T}$ bezeichnen wollen. Dabei wollen wir nicht nur ein einziges $(n+2)$-tupel (22) in Betracht ziehen, sondern ein einparametriges offenes Kontinuum $\mathfrak{C}$ von solchen Wertsystemen, das wir also in der Form

$$\mathfrak{C}: \quad \hat{T} = \hat{T}(\lambda), \ \hat{\gamma} = \hat{\gamma}(\lambda), \ \hat{\xi}_1^0 = \hat{\xi}_1^0(\lambda), \ \ldots, \ \hat{\xi}_n^0 = \hat{\xi}_n^0(\lambda), \tag{31}$$

$$\bar{\lambda} < \lambda < \bar{\bar{\lambda}} \tag{32}$$

ansetzen können, wobei λ der Scharparameter ist, der das Intervall (32) durchläuft, in welchem die $n+2$ Funktionen (31) stetig sein sollen. Man kann $\mathfrak{C}$ in einen Phasenraum wie folgt einbetten. Es bezeichne $\mathfrak{R}$ die Mannigfaltigkeit aller $(n+2)$-tupel von reellen oder komplexen Zahlen

$$\Theta, \gamma, \xi_1^0, \ldots, \xi_n^0 \tag{33}$$

[d. h. den „euklidischen" Raum von $n+2$ komplexen Dimensionen] und $\mathfrak{R}$ die Mannigfaltigkeit aller $(n+2)$-tupel von reellen Zahlen (33) [also den $(n+2)$-dimensionalen reellen euklidischen Raum, der ein Unterraum von $\mathfrak{R}$ ist], endlich $\mathfrak{G}$ die Mannigfaltigkeit aller der Bedingung (7) genügenden Punkte von $\mathfrak{R}$, die also, sofern (5) nicht identisch verschwindet, eine Überfläche in $\mathfrak{R}$ bildet (die auch nicht auf $\mathfrak{R}$ liegende Punkte enthält). Die Mannigfaltigkeit $\mathfrak{C}$ ist dann eine auf $\mathfrak{G}$ und zugleich auf $\mathfrak{R}$ verlaufende Kurve in $\mathfrak{R}$ mit der Parameterdarstellung

$$\begin{gathered}\mathfrak{C}: \quad \Theta = \hat{T}(\lambda), \ \gamma = \hat{\gamma}(\lambda), \ \xi_1^0 = \hat{\xi}_1^0(\lambda), \ \ldots, \ \xi_n^0 = \hat{\xi}_n^0(\lambda); \\ \bar{\lambda} < \lambda < \bar{\bar{\lambda}}.\end{gathered} \tag{34}$$

Daß es sich bei jedem λ um eine periodische und die invariante Relation (6) befriedigende Lösung handelt, können wir wie folgt formulieren. — Wie bei (24) und (26) betont wurde, hängen die Potenzreihen $F_0; F_1, \ldots, F_n$ nicht nur von ihren $n+2$ angeschriebenen Argumenten, sondern auch von den $n+2$ Bestimmungsstücken (22), also diesmal mit Rücksicht auf (34) einfach von λ ab, so daß wir an Stelle von (24), (26)

und (14) ausführlich

$$F_k = F_k^{[\lambda]} = F_k^{[\lambda]}(\vartheta, \Delta\gamma, \Delta\xi_1^0, \ldots, \Delta\xi_n^0); \quad k = 0, 1, \ldots, n; \quad \underline{\lambda} < \lambda < \bar{\lambda}, \tag{35}$$

$$\vartheta = \Theta - \hat{T}(\lambda), \quad \Delta\gamma = \gamma - \hat{\gamma}(\lambda), \quad \Delta\xi_1^0 = \xi_1^0 - \hat{\xi}_1^0(\lambda), \quad \ldots, \quad \Delta\xi_n^0 = \xi_n^0 - \hat{\xi}_n^0(\lambda) \tag{36}$$

schreiben wollen, wobei

$$\vartheta, \ \Delta\gamma, \ \Delta\xi_1^0, \ \ldots, \ \Delta\xi_n^0 \tag{37}$$

$n+2$ unabhängige (komplexe) Veränderliche sind. Daß die Lösung (10) für $\gamma = \hat{\gamma}(\lambda)$, $\xi_j^0 = \hat{\xi}_j^0(\lambda)$; $j = 1, \ldots, n$ auf $\mathfrak{G}$ liegt und die Periode $\hat{T}(\lambda)$ hat, kann dann mit Rücksicht auf das bei (27) und (25) Gesagte einfach durch

$$F_k^{[\lambda]}(0, 0, 0, \ldots, 0) = 0; \quad k = 0, 1, \ldots, n; \quad \underline{\lambda} < \lambda < \bar{\lambda} \tag{38}$$

ausgedrückt werden.

Jedes Wertsystem (9) legt eine Lösung (10) von (2) fest und umgekehrt. Jedem Punkte $P = P_\lambda$ der Kurve $\mathfrak{C}$:

$$P_\lambda = P(\hat{T}(\lambda), \hat{\gamma}(\lambda), \hat{\xi}_1^0(\lambda), \ldots, \hat{\xi}_n^0(\lambda)) \tag{39}$$

entspricht also eine der Bedingung (6) genügende, reelle, nach $\hat{T}(\lambda)$ periodische Lösung $\mathfrak{X}_\lambda$:

$$\mathfrak{X}_\lambda: \quad \hat{\xi}_i(\tau, \lambda) = \xi_i(\tau; \hat{\gamma}(\lambda); \hat{\xi}_1^0(\lambda), \ldots, \hat{\xi}_n^0(\lambda)) \qquad (i = 1, \ldots, n) \tag{40}$$

und umgekehrt, so daß, wenn $\mathfrak{M}$ die Mannigfaltigkeit aller Lösungen $\mathfrak{X}_\lambda$ bedeutet:

$$\mathfrak{M}: \quad \mathfrak{X}_\lambda, \quad \underline{\lambda} < \lambda < \bar{\lambda}, \tag{41}$$

$\mathfrak{M}$ und $\mathfrak{C}$ durch λ aufeinander topologisch bezogen werden.

Bei Anwendung auf das Strömgrensche Problem (§ 6) wird (2) das System der Thieleschen Differentialgleichungen ($n = 4$) und (6) das Jacobische Integral, γ die Jacobische Konstante bedeuten, während dem sogenannten Massenprozentsatz μ irgendein *fest* gewählter Wert zwischen 0 und 1 (z. B. bei Darwin $\mu = \frac{1}{11}$ bzw. $\mu = \frac{10}{11}$, in den Kopenhagener Rechnungen $\mu = \frac{1}{2}$) zu erteilen ist, so daß μ in den Funktionszeichen G_i, Γ steckt. $\mathfrak{M}$ ist also ein Stück einer Gruppe von periodischen Lösungen. Nun ist aber in den numerischen Untersuchungen eben die Jacobische Konstante, also $\gamma = \hat{\gamma}$. der Scharparameter von $\mathfrak{M}$ (nach welchem das analytische Fortsetzen erfolgt). Dies besagt, daß unsere Funktion

$$\hat{\gamma} = \hat{\gamma}(\lambda) \tag{42}$$

zu invertieren, d. h. λ als eine Funktion

$$\lambda = \lambda(\hat{\gamma}) \tag{43}$$

von $\hat{\gamma}$ auszudrücken und dann (43) in (31) einzutragen ist, wodurch die Lösung (40), d. h. (die Periode und) die zu $\tau = 0$ gehörigen Anfangswerte Funktionen von $\hat{\gamma}$ werden:

$$(44) \qquad \hat{T} = \hat{T}(\hat{\gamma}), \quad \hat{\xi}_1^0 = \hat{\xi}_1^0(\hat{\gamma}), \quad \ldots, \quad \hat{\xi}_n^0 = \hat{\xi}_n^0(\hat{\gamma}),$$

und daher $\hat{\gamma}$ als der Scharparameter des vorgelegten Stückes $\mathfrak{M}$ der vorgelegten Bahngruppe erscheint. Es ist dabei von einem *hinreichend kleinen* Stücke der Gruppe (etwa *ganz* in der Nähe der Librationspunkte oder der Massen oder *ganz* weit weg von diesen) von vornherein bekannt (vgl. den Anfang von § 3), daß die $n+1$ Funktionen (44) analytische Funktionen der Jacobischen Konstante $\hat{\gamma}$ sind. Im Kleinen wäre also, mit Ausnahme der Umgebung eventueller Verzweigungsstellen der Funktionen (44), die Einführung des Parameters ganz überflüssig, indem $\lambda = \hat{\gamma}$ die vorteilhafteste Wahl von λ wäre. Im Großen, d. h. für das ganze Gruppenstück $\mathfrak{M}$ (und erst recht für die volle Gruppe bis zu ihrem natürlichen Abschluß), können aber die Funktionen (44) mehrdeutige Funktionen von $\hat{\gamma}$ sein[10]), und sie werden im allgemeinen tatsächlich algebraische Verzweigungssingularitäten haben. [Durch die Existenz von solchen Verzweigungsstellen für $\hat{T} = \hat{T}(\lambda)$ wird erst eine scheinbare Diskrepanz behoben, die einerseits zwischen dem numerischen Material, andererseits zwischen einem Herglotzschen Satz über $\hat{T} = \hat{T}(\lambda)$ besteht[11]).] Die Berücksichtigung der Mehrdeutigkeit im Großen war ein springender Punkt bei den Erfolgen der Strömgrenschen numerischen Rechnungen, während man vorher (Darwin, Poincaré) Kontinua von periodischen Lösungen, die an sich zu einer und derselben Mannigfaltigkeit gehören, nicht als „Fortsetzungen" voneinander (der Mathematiker kann, wie wir beweisen werden, sagen: analytische Fortsetzungen voneinander), sondern als zwei getrennte Mannigfaltigkeiten auffaßte[10]), wodurch jede Systematik der periodischen Bahnen verhindert werden mußte: ohne Berücksichtigung der Mehrdeutigkeit der „Jacobischen" Darstellung (44) wäre auch das Prinzip des natürlichen Abschlusses, das wir beweisen wollen, bereits an Hand des Kopenhagener numerischen Materials widerlegbar. Die Einführung eines *im Großen uniformisierenden* Parameters λ ist also bei unseren Zwecken nicht zu vermeiden: die Funktion (42) ist nirgends konstant, aber sie wird nur *abteilungsweise* monoton und daher die inverse Funktion (43) *im Großen mehrdeutig* sein können.

---

[10]) Vgl. I; s. auch XII.

[11]) Vgl. G. Herglotz, Seeliger-Festschrift (1924), S. 197–199, bzw. X, S. 35 oder XII, S. 671.

§ 3.

E. Strömgrens dynamisches Abschlußprinzip.

Das Gruppenstück $\mathfrak{M}$ [vgl. (41)] sei gegeben. Wie man die Existenz eines solchen, eventuell sehr *kleinen* Stückes einer Gruppe — etwa ganz in der Nähe der Librationszentra oder der Massen oder ganz weit weg von diesen — analytisch beweisen kann, ist heute ein prinzipiell erledigtes Problem. Hingegen ist mathematisch so gut wie gar nichts bekannt über das Problem *im Großen*, worüber E. Strömgren[6]) das folgende dynamische Prinzip ausgesprochen und — vorläufig mit Ausnahme seiner Gruppe g — bei allen seinen Gruppen numerisch bestätigt hat: Das Gruppenstück $\mathfrak{M}$ kann nur einen natürlichen, nämlich dynamisch bedingten Abschluß erleiden, d. h. die Bahnen von $\mathfrak{M}$ können in ihrer Abhängigkeit von $\hat{\gamma}$ auf reelle Weise und der Bedingung (6) gemäß jedenfalls so lange analytisch fortgesetzt werden, bis entweder mindestens eine der Funktionen ξ_i, oder die Periode, oder endlich $\hat{\gamma}$ (d. i. die Energie) unendlich wird [insbesondere kann also eine feste Stoßbahn keinesfalls den natürlichen Abschluß bilden, eine solche ist nämlich in den zugrunde gelegten Regularisierungskoordinaten den Nichtstoßbahnen gegenüber gar nicht ausgezeichnet [vgl. S. 323 und § 6]. Ich will hier die Richtigkeit dieses, das Verhalten im Großen regulierenden Strömgrenschen Prinzips analytisch *beweisen*. Es folgt dann einerseits die mathematische Sicherstellung der bisherigen numerischen Ergebnisse, andererseits die Existenz eines zur Zeit numerisch nicht erreichten natürlichen Abschlusses auch für die Gruppe g, wobei freilich auch die Möglichkeit besteht, daß die Gruppe „in sich geschlossen" ist — vgl. weiter unten —, so daß die erwähnten Bestimmungsstücke bei unbegrenzter analytischer Fortsetzung überhaupt beschränkt bleiben. — Freilich ist dabei auch die unbegrenzte Annäherung an eine der auf S. 328 ausgeschlossenen (fünf) Gleichgewichtslösungen als ein natürlicher Abschluß aufzufassen.

Es handelt sich also u. a. um folgendes. Unter $r = r(\lambda)$ bei gegebenem λ eine für $-\infty < \tau < +\infty$ gültige Schranke der n reellen Funktionen (40), also z. B. die Zahl

$$r(\lambda) = \max_{0 \leqq \tau \leqq \hat{T}(\lambda)} \sum_{i=1}^{n} [\hat{\xi}_i(\tau; \lambda)]^2 \tag{45}$$

und unter $m = m(\lambda)$ eine Schranke der drei Zahlen $r(\lambda)$, $\hat{T}(\lambda)$, $\hat{\gamma}(\lambda)$, also z. B. die Zahl

$$m(\lambda) = r(\lambda) + |\hat{\gamma}(\lambda)| + \hat{T}(\lambda) \tag{46}$$

verstanden, sei in dem Intervall (32)

$$\liminf_{\lambda \to \bar{\bar{\lambda}} - 0} m(\lambda) < +\infty. \tag{47}$$

Dann kann das Gruppenstück $\mathfrak{M}$ über den oberen Grenzpunkt $\bar{\bar{\lambda}}$ von (32)

hinaus fortgesetzt werden (insbesondere muß dann also auch

$$(47)' \qquad \limsup_{\lambda \to \bar{\bar{\lambda}} - 0} m(\lambda) < +\infty$$

gelten). Die Möglichkeit einer Fortsetzung soll hier zunächst folgendes heißen: Einerseits gibt es eine Zahl $\bar{\bar{\lambda}}' > \bar{\bar{\lambda}}$ derart, daß die reellen Funktionen (31) nicht nur in dem Intervall (32), sondern auch in dem größeren Gebiete

$$(48) \qquad \bar{\lambda} < \lambda < \bar{\bar{\lambda}}' \qquad (\bar{\bar{\lambda}}' > \bar{\bar{\lambda}})$$

stetig und derart sind, daß die Lösungen (40) nicht nur in (32), sondern auch in (48) reelle, nach $\hat{T}(\lambda)$ periodische Lösungen von (2) und (6) darstellen; andererseits entsteht das neu hinzukommende (nämlich zu $\bar{\bar{\lambda}} - 0 \leqq \lambda < \bar{\bar{\lambda}}'$ gehörige) Gruppenstück (wie klein dabei auch $\bar{\bar{\lambda}}' - \bar{\bar{\lambda}} > 0$ sein mag) aus dem gegebenen (nämlich zu $\bar{\lambda} < \lambda < \bar{\bar{\lambda}}$ gehörigen) Gruppenstück $\mathfrak{M}$ *nicht* einfach dadurch, daß man eine linksseitige Umgebung $\bar{\bar{\lambda}}'' < \lambda < \bar{\bar{\lambda}}$ von $\bar{\bar{\lambda}}$ rückläufig, d. h. von $\bar{\bar{\lambda}}$ bis $\bar{\bar{\lambda}}''$ durchläuft ($\bar{\lambda} < \bar{\bar{\lambda}}'' < \bar{\bar{\lambda}} < \bar{\bar{\lambda}}'$), da doch sonst, d. h. im Falle einer solchen „Spiegelung", die zu $\lambda = \bar{\bar{\lambda}}$ gehörige Bahn als eine (dynamisch unverständliche) *Grenz*bahn anzusehen und damit das Prinzip des natürlichen (d. h. dynamisch bedingten) Abschlusses gerade in dem — wie eben unser Satz lehrt, stets vorliegenden! — Falle, daß die zu λ gehörige Lösung bei dem Grenzübergang $\lambda \to \bar{\bar{\lambda}} - 0$ gegen eine periodische Bahn konvergiert, ebenso inhaltslos wäre, wie gewisse Aussagen von Poincaré über das Verschwinden von periodischen Bahnen es sind. Erst recht ist es daher die Poincarésche Aussage über Stabilitätsaustausch[11a]), nach welcher zwei Scharen von periodischen Lösungen[12]) nach ihrer Vereinigung nur dann zum Verschwinden kommen können, wenn, im Sinne der charakteristischen Exponenten, die eine stabil, die andere labil ist; daß übrigens auch der Beweis von vornherein unbefriedigend sein muß, da dabei der Verzweigungssatz zugrunde liegt, bei dem aber auch das „Problem im Großen" übersehen wird (vgl. S. 325 oben in der Einleitung), versteht sich von selbst. Diese Bemerkungen gelten auch für die Poincarésche Behandlung des Problems der Doppel- bzw. mehrfachen Bahnen[13]) (Kommensurabelwerden des charakteristischen Exponenten), doch will ich hier auf diese Problematik des Poincaréschen zweiten Genres nicht näher eingehen.

[11a]) XIV, § 378—380.

[12]) Man kann übrigens — zum mindesten ohne eine funktionentheoretische Charakterisierung im Kleinen, die aber für das dynamische Problem keine *direkte* Interpretation zuläßt — jeden, noch so regulären inneren Punkt eines Gruppenstückes als einen Punkt auffassen, von welchem „zwei" Zweige „abzweigen" (der eine nach rechts, der andere nach links).

[13]) Vgl. die auf S. 323 bei der zweiten Sorte und dem zweiten Genre angegebene Literatur, die zitierte (S. 322) sich auf Darwin beziehende Stelle in XIV, ferner II, S. 25—26 (group n) und Math. Annalen **96** (1926), S. 282 ff.

Aus unserem Satze folgt übrigens, daß ein Verschwinden der periodischen Lösung einfach unmöglich ist, so daß die Poincarésche Aussage, daß nämlich periodische Lösungen *höchstens* paarweise — aber, wie aus unserem Satze folgt, *in Wirklichkeit überhaupt nicht*! — verschwinden können, bereits aus diesem Grunde inhaltslos ist; oben handelt es sich aber darum, daß die Poincarésche Aussage über das (höchstens) *paarweise* Verschwinden, wie ich bei einer früheren Gelegenheit schon ausgeführt habe, auch sonst trivial wäre, indem man einfach durch die oben erwähnte „Spiegelung" von $\bar{\bar{\lambda}}'' < \lambda < \bar{\bar{\lambda}}$ in $\bar{\bar{\lambda}} < \lambda < \bar{\bar{\lambda}}'$ erreichen kann, daß aus $\bar{\bar{\lambda}}$ ein *Paar* von linearen Scharen ausgeht, wie regulär die Stelle $\bar{\lambda}$ auch sein mag; eben diese unhaltbare Terminologie hat zu den oben (S. 333) erwähnten analytisch durchaus falschen Klassifikationen geführt; Darwin nimmt bei seiner — heute unhaltbaren — Einteilung in families ausdrücklich auf den Poincaréschen Satz Bezug. — Übrigens ist bei Poincaré auch der Beweis seiner *höchstens im Sinne der Puiseuxschen Sätze, aber dynamisch keinesfalls sinnvollen* Aussage fehlerhaft (s. S. 335 oben), und die vorliegende Arbeit gilt eigentlich eben einer analytischen Legalisierung des Poincaréschen Ansatzes, freilich dann nicht in der Poincaréschen, sondern in der ganz anderen Richtung, welche durch die in § 2 erwähnte grundlegende Strömgrensche Bemerkung, *die eine inhaltsvolle Fragestellung überhaupt ermöglicht*, für den Mathematiker eindeutig vorgezeichnet wird. Daß es sich bei der in dem vorigen Absatz festgelegten Fortsetzung mit Rücksicht auf (42), (43), (44) um eine analytische Fortsetzung handelt, wird aus dem Beweis hervorgehen. — Wir haben uns bisher auf den oberen Endpunkt von (32) beschränkt und wollen das auch künftig tun; bei dem unteren Endpunkt $\bar{\lambda}$ von (32) verfährt man nämlich analog.

Um eine triviale, d. h. in dem obigen Sinne „gespiegelte" Fortsetzung auszuschließen, haben wir uns nur *mit gegebenen Gruppenstücken* $\mathfrak{M}$ zu beschäftigen, *die keine „gespiegelten" Teilstücke enthalten*; die Behauptung ist dann, daß ein derart gewähltes Gruppenstück durch unbegrenzte analytische Fortsetzung (und zwar ohne Benutzung von gespiegelten Fortsetzungen) entweder sich in sich schließt oder eine Gleichgewichtslage erreicht oder die Funktion $m(\lambda)$ gegen $+\infty$ gehen läßt (oder mehrere dieser Möglichkeiten eintreten), wobei zwei Lösungen dann[14]) und nur

[14]) Zu beachten ist dabei die Willkürlichkeit der Epoche $\tau = 0$ in (4); vgl. auch die Definition der Mannigfaltigkeit $g = 0$ weiter unten (S. 344). — Es handelt sich darum, daß etwa die beiden Anfangswertsysteme

$$-\xi_1^0, \xi_2^0, \ldots, \xi_n^0; \quad \xi_1^0, \xi_2^0, \ldots, \xi_n^0$$

an und für sich eine und dieselbe Bahnkurve definieren können. Darin ist insbesondere die Möglichkeit mit einbegriffen, daß *eine Doppelbahn einen Wendepunkt der Gruppenmannigfaltigkeit bildet.*

dann als *identisch* gelten, wenn sie durch ein und dasselbe $(n+2)$-tupel (22) charakterisiert sind, während die nähere analytische Definition des Insichgeschlossenseins bis S. 347 verschoben werden soll.

§ 4.

Beispiele zum Abschlußprinzip.

Es ist oben (S. 334) schon erwähnt worden, daß die Funktion (46) bei gewissen Gruppen beschränkt bleiben kann, wie weit man auch die Fortsetzung der Gruppe treiben will. Dies wird der Fall sein, wenn die Gruppe in sich geschlossen ist, d. h. wenn ihre Bahnen auf die Punkte der Achse $-\infty < \lambda < +\infty$ derart bezogen werden können, daß bei jedem λ zu λ und zu $\lambda+1$ dieselbe Bahn gehört; solche geschlossene Gruppen hat E. Strömgren ebenfalls numerisch gefunden. Daß eine Gruppe geschlossen ist, läßt sich auf Grund unseres Satzes oft *von vornherein* beweisen.

Es ist leicht zu zeigen, daß bei den drei kollinearen Librationspunkten des restringierten Dreikörperproblems (Gruppen a, b, c), und zwar bei jedem Werte des Massenprozentsatzes, von der Gleichgewichtslösung, der etwa der Parameterwert $\lambda = \bar{\lambda}$ entsprechen möge, nur eine einzige Schar von periodischen Lösungen ausgeht, die für $\lambda \to \bar{\lambda} + 0$ gegen die Gleichgewichtslösung konvergiert [die Periode $\hat{T}(\lambda)$ konvergiert dabei gegen einen wohlbestimmten Grenzwert > 0, der durch das 2π-fache des reziproken Wertes des Betrages des zu $\lambda = \bar{\lambda}$ gehörigen stabilen Exponenten — der andere ist labil — gegeben ist; daß für $\lambda \to \bar{\lambda} + 0$ die Jacobische Konstante $\hat{\gamma}(\lambda)$ gegen einen endlichen Grenzwert konvergiert, ist trivial]. Unter Benutzung dieses Umstandes können wir mit Rücksicht auf das Abschlußprinzip manches über den Entwicklungsgang der kollinearen Librationsgruppen aussagen, doch kann darauf erst im Rahmen der Librationsgeneagogie eingegangen werden, wobei auch die Vergleichung mit dem numerischen Material durchgeführt werden soll.

Das Abschlußprinzip besagt nur, daß eine Fortsetzung der Gruppe jedenfalls so lange möglich ist, als die Funktion (46) nicht unendlich wird, es behauptet aber nicht, daß eine analytische Fortsetzung über eine Unendlichkeitsstelle der Funktion (46) hindurch unmöglich sein muß. Die Sache liegt z. B. bei den (direkten bzw. retrograden) massennahen Gruppen f und g so, daß sie, eben in ihren massennahen Gruppenstücken, analytische Fortsetzungen *voneinander* sind (sie brauchen also, trotz der eindeutigen Bestimmtheit ihrer massennahen Gruppenstücke, bereits aus diesem Grunde nicht geschlossen zu sein), indem die Potenzreihen für die Fourierkoeffizienten der periodischen Lösungen nach dem reziproken Werte der $\frac{1}{2}$-ten Potenz der Jacobischen Konstante, d. h. von $\hat{\gamma}$ fortschreiten und für sehr

große Werte von $\hat{\gamma}$, d. h. eben für das massennahe Gruppenstück der Gruppe f ($\sqrt{\hat{\gamma}} < 0$) und der Gruppe g ($\sqrt{\hat{\gamma}} > 0$), also in der Umgebung von $\sqrt{\hat{\gamma}} = \pm\infty$, d. h. von $\hat{\gamma}^{-\frac{1}{2}} = \pm 0$, konvergent sind und daher in der Umgebung der Stelle $\hat{\gamma} = +\infty$ für $\sqrt{\hat{\gamma}} > 0$ und für $\sqrt{\hat{\gamma}} < 0$ (d. h. für g und f) eine und dieselbe analytische Funktion darstellen müssen. — Durchaus analog liegen die Verhältnisse bei dem Gruppenpaar l, m, nur daß es sich diesmal um die beiden massenfernen Gruppenstücke handelt, indem an der Stelle, an welcher sie analytische Fortsetzungen voneinander sind, nicht nur die Jacobische Konstante, sondern auch die Periode bzw. die Bahndimension unendlich wird.

Diese Betrachtungen ließen sich vertiefen und die Beispiele vervielfältigen, insbesondere durch Einbeziehung der Gruppe k und vor allem der dem Mathematiker bis zur letzten Zeit so sehr mysteriösen „nichtinfinitesimalen" äquilateralen Librationsgruppen[6]), [15]) (bei nicht kleinem Massenprozentsatz) und ihres Zusammenhanges mit den beiden (nur bei kleinem Massenprozentsatz existierenden) „infinitesimalen" Librationsgruppen um einen äquilateralen Librationspunkt [es handelt sich[15]) auch bei nicht kleinem Prozentsatz um die Gruppen d, e; vgl. die Schlußbemerkung des vorliegenden Aufsatzes]. Doch soll darauf in einer späteren Mitteilung im Rahmen einer ins einzelne gehenden mathematischen Systematik aller Gruppen eingegangen werden, während es hier nur darauf ankam, die den Endverlauf regulierende Kraft des Abschlußsatzes an Hand von möglichst einfachen Beispielen zu illustrieren.

Es sei noch erwähnt, daß es an sich denkbar wäre, daß das *Kontinuum*, welches aus den periodischen Bahnen einer Gruppe besteht, eine *Verzweigungsstelle* hat, so daß aus einem inneren Punkte des Kontinuums nicht nur zwei, sondern mindestens drei lineare Scharen ausgehen (deren zwei gibt es bei jedem inneren Punkt), wie dies z. B. bei dem Abzweigen der Jacobischen Flüssigkeitsellipsoide von den Maclaurinschen der Fall ist. doch ist bisher keine einzige Gruppe bekannt, in deren Entwicklung die Existenz eines solchen „Stadiums des Indeterminismus" nachweisbar wäre, so daß die im Rahmen der Anschmiegungsmethoden leider schon eingebürgerte Ausdrucksweise „Verzweigungsgleichung" geradezu irreführend ist.

Wenn $\lambda = \bar{\bar{\lambda}}$ eine Stelle des Indeterminismus ist, so hat man sich bei der Fortsetzung über $\lambda = \bar{\bar{\lambda}}$ hinaus für *irgendeinen* der zur Verfügung stehenden, von dem zu $\lambda = \bar{\bar{\lambda}}$ gehörigen Zweig verschiedenen Äste zu entscheiden. Die Ausführungen von § 3 sind in diesem Sinne zu verstehen. Dies wird freilich auch aus dem Beweise des Abschlußsatzes hervorgehen.

---

[15]) Vgl. XI, § 4.

§ 5.

Beweis des Abschlußsatzes mittels analytischer Fortsetzung.

Wir gehen nun zum Beweise des in § 3 analytisch formulierten Strömgrenschen Abschlußprinzips über.

Nach Voraussetzung gilt (47), d. h. es gibt eine Zahl B und dazu in (32) eine gegen $\bar{\bar{\lambda}}$ konvergierende Punktfolge $\{\lambda'_\nu\}$ mit

$$m(\lambda'_\nu) \leqq B \quad \text{für} \quad \nu = 0, 1, \ldots; \quad \lambda'_0 < \lambda'_1 < \ldots < \lambda'_\nu \to \bar{\bar{\lambda}} \tag{48}$$

(das Zeichen $\to$ soll sich immer auf den Grenzübergang $\nu \to +\infty$ beziehen). Da die $n+2$ Zahlenfolgen

$$\hat{T}(\lambda'_\nu),\ \hat{\gamma}(\lambda'_\nu),\ \hat{\xi}^0_1(\lambda'_\nu),\ \ldots,\ \hat{\xi}^0_n(\lambda'_\nu) \qquad (\nu = 0, 1, \ldots) \tag{49}$$

mit Rücksicht auf (43), (46), (45) und (4) beschränkt sind, so enthält die Folge $\{\lambda'_\nu\}$ eine Auswahlfolge $\{\lambda_\nu\}$ derart, daß die entsprechenden Teilfolgen

$$\hat{T}(\lambda_\nu),\ \hat{\gamma}(\lambda_\nu),\ \hat{\xi}^0_1(\lambda_\nu),\ \ldots,\ \hat{\xi}^0_n(\lambda_\nu) \qquad (\nu = 0, 1, \ldots) \tag{50}$$

der Folgen (49) gegen endliche Grenzwerte streben, die

$$\Theta^*,\ \gamma^*,\ \xi^*_1,\ \ldots,\ \xi^*_n \tag{51}$$

heißen mögen:

$$\begin{gathered}\hat{T}(\lambda_\nu) \to \Theta^*,\ \hat{\gamma}(\lambda_\nu) \to \gamma^*,\ \hat{\xi}^0_1(\lambda_\nu) \to \xi^*_1,\ \ldots,\ \hat{\xi}^0_n(\lambda_\nu) \to \xi^*_n; \\ \bar{\lambda} < \lambda_\nu < \lambda_{\nu+1} \to \bar{\bar{\lambda}}.\end{gathered} \tag{52}$$

Für spätere Zwecke wollen wir gleich hier die Abkürzungen

$$\begin{gathered}\vartheta_\nu = \hat{T}(\lambda_\nu) - \hat{T}(\lambda_0), \quad \Delta_\nu \gamma = \hat{\gamma}(\lambda_\nu) - \hat{\gamma}(\lambda_0), \\ \Delta_\nu \xi^0_1 = \hat{\xi}^0_1(\lambda_\nu) - \hat{\xi}^0_1(\lambda_0), \quad \ldots, \quad \Delta_\nu \xi^0_n = \hat{\xi}^0_n(\lambda_\nu) - \hat{\xi}^0_n(\lambda_0)\end{gathered} \tag{53}$$

und

$$\begin{gathered}\vartheta^* = \Theta^* - \hat{T}(\lambda_0), \quad \Delta^* \gamma = \gamma^* - \hat{\gamma}(\lambda_0), \\ \Delta^* \xi_1 = \xi^*_1 - \hat{\xi}^0_1(\lambda_0), \quad \ldots, \quad \Delta^* \xi_n = \xi^*_n - \hat{\xi}^0_n(\lambda_0)\end{gathered} \tag{54}$$

einführen (wobei die Abweichungen (53) für $\nu = 0$ offenbar verschwinden). Da $\{\lambda_\nu\}$ eine Teilfolge von $\{\lambda'_\nu\}$ ist, so gilt wegen (48)

$$m(\lambda_\nu) \leqq B, \quad \nu = 0, 1, \ldots; \quad \lambda_0 < \lambda_1 < \ldots < \lambda_\nu \to \bar{\bar{\lambda}}, \tag{55}$$

so daß wegen (46), (45) und (4) alle unter (50) angegebenen Zahlen dem Betrage nach $\leqq B$ sind. Wählen wir also irgendwie zwei feste Zahlen L, M, die $\geqq B$ sind, so wird einerseits (13) für $\hat{T} = \hat{T}(\lambda_\nu)$ und andererseits wegen (54), (46), (45) und mit Rücksicht auf die Periodizität der Funktionen (40) — die Ungleichung

$$\max_{-\infty < \tau < +\infty} |\hat{\xi}(\tau, \lambda_\nu)| \leqq M$$

und daher auch die zu (11) gehörige Bedingung (12) gleichmäßig für alle ν erfüllt sein, so daß wir, wie auf S. 329, Z. 9 betont wurde, bei jedem fest gewählten $\varepsilon > 0$ die Zahl $\delta = \delta_\varepsilon > 0$ des Existenzsatzes von § 1 unabhängig von ν wählen können. Dies besagt u. a., daß die Potenzreihen

$$F_k^{[\lambda_\nu]} = F_k^{[\lambda_\nu]}(\vartheta, \Delta\gamma, \Delta\xi_1^0, \ldots, \Delta\xi_n^0) \quad (k = 0, 1, \ldots, n), \tag{56}$$

wobei an Stelle der Schreibweise (21), (27) die bei (38) eingeführte Bezeichnung benutzt ist, in dem komplexen Bereiche (19) ihrer unabhängigen Argumente

$$\begin{gathered} \vartheta = \Theta - \hat{T}(\lambda_\nu), \quad \Delta\gamma = \gamma - \hat{\gamma}(\lambda_\nu), \\ \Delta\xi_1^0 = \xi_1^0 - \hat{\xi}_1^0(\lambda_\nu), \quad \ldots, \quad \Delta\xi_n^0 = \xi_n^0 - \hat{\xi}_n^0(\lambda_\nu) \end{gathered} \tag{57}$$

(wobei also, bei festem ν, die Koordinaten (33) des komplexen Raumes $\mathfrak{R}$ als unabhängige Veränderliche zugrunde liegen) regulär sein müssen.

Nach erfolgter Wahl von ε und δ_ε können wir mit Rücksicht auf (52) eine feste ganze Zahl N finden derart, daß für alle $\nu \geqq N$

$$\begin{gathered} |\hat{T}(\lambda_\nu) - \Theta^*| \leqq \tfrac{\varepsilon}{4}, \quad |\hat{\gamma}(\lambda_\nu) - \gamma^*| \leqq \tfrac{\delta_\varepsilon}{4}, \\ |\hat{\xi}_1^0(\lambda_\nu) - \xi_1^*| \leqq \tfrac{\delta_\varepsilon}{4}, \quad \ldots, \quad |\hat{\xi}_n^0(\lambda_\nu) - \xi_n^*| \leqq \tfrac{\delta_\varepsilon}{4} \end{gathered} \tag{58}$$

gilt und wir können voraussetzen, daß (58) bereits von $\nu = 0$ an richtig ist; denn wäre $N \neq 0$, so könnte $N + \nu = \nu'$ als neuer Zeiger eingeführt und sodann ν' einfach durch ν bezeichnet werden. — Da so (58) für $\nu = 0, 1, \ldots$ gilt, so folgt für die Zahlen (53) bei jedem $\nu \geqq 0$

$$|\vartheta_\nu| \leqq \tfrac{\varepsilon}{2}, \quad |\Delta_\nu\gamma| \leqq \tfrac{\delta_\varepsilon}{2}, \quad |\Delta_\nu\xi_1^0| \leqq \tfrac{\delta_\varepsilon}{2}, \quad \ldots, \quad |\Delta_\nu\xi_n^0| \leqq \tfrac{\delta_\varepsilon}{2}, \tag{59}$$

also — da doch nach (52), (53), (54)

$$\vartheta_\nu \to \vartheta^*, \quad \Delta_\nu\gamma \to \Delta^*\gamma, \quad \Delta_\nu\xi_1^0 \to \Delta^*\xi_1, \quad \ldots, \quad \Delta_\nu\xi_n^0 \to \Delta^*\xi_n \tag{60}$$

gilt — offenbar auch

$$|\vartheta^*| \leqq \tfrac{\varepsilon}{2}, \quad |\Delta^*\gamma| \leqq \tfrac{\delta_\varepsilon}{2}, \quad |\Delta^*\xi_1| \leqq \tfrac{\delta_\varepsilon}{2}, \quad \ldots, \quad |\Delta^*\xi_n| \leqq \tfrac{\delta_\varepsilon}{2}. \tag{61}$$

Nun sind aber die $n+1$ Funktionen

$$F_k^{[\lambda_0]} = F_k^{[\lambda_0]}(\vartheta, \Delta\gamma, \Delta\xi_1^0, \ldots, \Delta\xi_n^0) \quad (k = 0, 1, \ldots, n) \tag{62}$$

in dem komplexen Bereiche (19), wobei

$$\begin{gathered} \vartheta = \Theta - \hat{T}(\lambda_0), \quad \Delta\gamma = \gamma - \hat{\gamma}(\lambda_0), \\ \Delta\xi_1^0 = \xi_1^0 - \hat{\xi}_1^0(\lambda_0), \quad \ldots, \quad \Delta\xi_n^0 = \xi_n^0 - \hat{\xi}_n^0(\lambda_0) \end{gathered} \tag{63}$$

zu setzen ist und die komplexen Koordinaten (33) des Raumes $\mathfrak{R}$ die unabhängigen Veränderlichen sind, nach dem Schluß des vorigen Absatzes in dem Bereiche (19) gewiß regulär, so daß (53) und (54) wegen (59) und

(61) in dem Regularitätsbereiche (19) von (62) liegen. Andererseits ist sobald (63) die Bedingung (19) befriedigt — nach den bei (24), (25) und (26) gemachten Bemerkungen das Erfülltsein der $n+1$ Bedingungsgleichungen

$$F_k^{[\lambda_0]}(\vartheta, \Delta\gamma, \Delta\xi_1^0, \ldots, \Delta\xi_n^0) = 0 \qquad (k = 0, 1, \ldots, n) \tag{64}$$

ganz allgemein die *notwendige und hinreichende* Bedingung dafür, daß die durch das [(3) entsprechende] Wertsystem

$$\gamma = \hat{\gamma}(\lambda_0) + \Delta\gamma, \qquad \xi_1^0 = \hat{\xi}_1^0(\lambda_0) + \Delta\xi_1^0, \ldots, \xi_n^0 = \hat{\xi}_n^0(\lambda_0) + \Delta\xi_n^0 \tag{65}$$

festgelegte Lösung (8) von (2) periodisch ist, die Periode

$$\Theta = \hat{T}(\lambda_0) + \vartheta \tag{66}$$

hat und auf der invarianten Mannigfaltigkeit (6) verläuft. Aus der *Notwendigkeit* dieser Bedingung folgt zunächst bei jedem ν, daß die Größen (53) die Gleichungen

$$F_k^{[\lambda_0]}(\vartheta_\nu, \Delta_\nu\gamma, \Delta_\nu\xi_1^0, \ldots, \Delta_\nu\xi_n^0) = 0 \qquad (k = 0, 1, \ldots, n) \tag{67}$$

befriedigen müssen, da doch $\lambda = \lambda_\nu$ in dem Parametergebiete (32) des gegebenen Gruppenstückes $\mathfrak{M}$ liegt, und aus (67) folgt wegen (60) für $\nu \to +\infty$

$$F_k^{[\lambda_0]}(\vartheta^*, \Delta^*\gamma, \Delta^*\xi_1, \ldots, \Delta^*\xi_n) = 0 \qquad (k = 0, 1, \ldots, n) \tag{68}$$

und wegen (53) für $\nu = 0$

$$F_k^{[\lambda_0]}(0, 0, 0, \ldots, 0) = 0, \qquad (k = 0, 1, \ldots, n), \tag{69}$$

wie es schon unter (38) angegeben wurde. Es ist aber vor allem die *Hinlänglichkeit* der Bedingung (64) von Bedeutung, und es wird dabei bequem sein, das notwendige und hinreichende Kriterium einer identischen Umformung zu unterwerfen, und zwar wie folgt.

Führt man in (62) an Stelle der in (62) angeschriebenen Argumente (63) mittels

$$\begin{gathered} \nabla\vartheta = \vartheta - \vartheta^*, \qquad \nabla\gamma = \Delta\gamma - \Delta^*\gamma, \\ \nabla\xi_1^0 = \Delta\xi_1^0 - \Delta^*\xi_1, \quad \ldots, \quad \nabla\xi_n^0 = \Delta\xi_n^0 - \Delta^*\xi_n \end{gathered} \tag{70}$$

neue unabhängige Variable

$$\nabla\vartheta, \quad \nabla\gamma, \quad \nabla\xi_1^0, \ldots, \nabla\xi_n^0 \tag{70'}$$

ein, indem man die Potenzreihen (62) nach Potenzen der Veränderlichen (70) ordnet, d. h. die Funktionen (62) vermöge (70) identisch gleich den Funktionen

$$F_k^V = F_k^V(\nabla\vartheta, \nabla\gamma, \nabla\xi_1^0, \ldots, \nabla\xi_n^0) \qquad (k = 0, 1, \ldots, n) \tag{71}$$

setzt, so sind die Potenzreihen (71) wegen (19), (61) und (70) in dem Bereiche

$$|\nabla\vartheta| \leqq \frac{\varepsilon}{2}, \quad |\nabla\gamma| \leqq \frac{\delta_\varepsilon}{2}, \quad |\nabla\xi_1^0| \leqq \frac{\delta_\varepsilon}{2}, \quad \ldots, \quad |\nabla\xi_n^0| \leqq \frac{\delta_\varepsilon}{2} \tag{72}$$

offenbar regulär. Die noch zu dem gegebenen Gruppenstück $\mathfrak{M}$ gehörigen Wertsysteme

$$\begin{gathered} \nabla_\nu\vartheta = \vartheta_\nu - \vartheta^*, \quad \nabla_\nu\gamma = \Delta_\nu\gamma - \Delta^*\gamma, \\ \nabla_\nu\xi_1^0 = \Delta_\nu\xi_1^0 - \Delta^*\xi_1, \ldots, \nabla_\nu\xi_n^0 = \Delta_\nu\xi_n^0 - \Delta^*\xi_n \end{gathered} \tag{73}$$

genügen wegen (68) der Bedingung

$$\nabla_\nu\vartheta \to 0, \quad \nabla_\nu\gamma \to 0, \quad \nabla_\nu\xi_1^0 \to 0, \quad \ldots, \quad \nabla_\nu\xi_n^0 \to 0 \tag{74}$$

und sie liegen wegen (53), (54), (58) bei jedem ν in dem Regularitätsbereich (72). — Ist einmal die Bedingung (72) erfüllt, so ist das Erfülltsein der $n+1$ Bedingungsgleichungen

$$F_k^V(\nabla\vartheta, \nabla\gamma, \nabla\xi_1^0, \ldots, \nabla\xi_n^0) = 0 \qquad (k = 0, 1, \ldots, n) \tag{75}$$

für die $n+2$ Zuwüchse die notwendige und hinreichende Bedingung dafür, daß die durch das [(3) entsprechende] Wertsystem

$$\begin{gathered} \gamma = \hat{\gamma}(\lambda_0) + \Delta^*\gamma + \nabla\gamma, \\ \xi_1^0 = \hat{\xi}_1^0(\lambda_0) + \Delta^*\xi_1 + \nabla\xi_1^0, \quad \ldots, \quad \xi_n^0 = \hat{\xi}_n^0(\lambda_0) + \Delta^*\xi_n + \nabla\xi_n^0 \end{gathered} \tag{76}$$

festgelegte Lösung (8) von (2) periodisch ist, die Periode

$$\Theta = \hat{T}(\lambda_0) + \vartheta^* + \nabla\vartheta \tag{77}$$

hat und auf der invarianten Mannigfaltigkeit (8) verläuft; es handelt sich ja um eine identische Umformung von (68). — Insbesondere die Beziehungen (67) bzw. (69) nehmen jetzt die Gestalt

$$F_k^V(\nabla_\nu\vartheta, \nabla_\nu\gamma, \nabla_\nu\xi_1^0, \ldots, \nabla_\nu\xi_n^0) = 0 \qquad (k = 0, 1, \ldots, n) \tag{78}$$

bzw.

$$F_k^V(0, 0, 0, \ldots, 0) = 0 \qquad (k = 0, 1, \ldots, n) \tag{79}$$

an. Die Jacobische Determinante

$$\frac{\partial(F_0^V, F_1^V, \ldots, F_n^V)}{\partial(\nabla\vartheta, \nabla\xi_1^0, \ldots, \nabla\xi_n^0)},$$

die auch den Parameter $\nabla\gamma$ enthält, kann freilich verschwinden (vgl. weiter unten).

Das Verschwinden einer Determinante ist bekanntlich die Grundschwierigkeit der Poincaréschen Problematik. Man beachte aber, daß das Problem bei Poincaré stets auf eine überbestimmte und im allgemeinen sogar unlösbare Weise gestellt ist, indem Poincaré entweder über die Periodenstörung ϑ (XIII, S. 79ff., auch S. 83, ferner S. 139ff.), oder aber über einen ξ-Zuwachs (z. B. XIII, S. 90) auf eine durchaus unnatürliche Weise willkürlich disponiert; was aber bei Poincaré diesbezüglich vor allem fehlt, ist der Gruppenbegriff bzw. die grundsätzliche Unterscheidung des Fortsetzungsproblems von dem Problem der isoliert gegebenen Bahnen,

d. h. von dem Anschmiegungsproblem; vgl. hierzu XII, 2. Absatz v. u. — Für unseren Beweis ist die Beachtung aller dieser Umstände sehr wesentlich, sonst würde man nämlich auf dieselben Schwierigkeiten stoßen, an welchen die Behandlung derartiger Probleme bei Poincaré bereits im Kleinen scheitern mußte.

* * *

Es kann angenommen werden, daß die abzählbar vielen Lösungen (73) und (75) voneinander durchweg (d. h. nicht nur der Epoche nach) verschiedene Bahnkurven $\xi_i = \xi_i(\tau)$ festlegen. Denn ließe sich dies bei keiner Wahl der gegen $\bar{\lambda}$ konvergierenden Folge $\{\lambda_\nu\}$ erreichen, so würde das Gruppenstück $\mathfrak{M}$ überhaupt nur eine einzige periodische Lösung enthalten, wäre also gar nicht ein Gruppenstück (es ist bei unserer Fragestellung der springende Punkt, daß sie die Frage nach der analytischen Fortsetzbarkeit eines gegebenen Kontinuums periodischer Lösungen von dem Problem der isolierten periodischen Lösungen wesentlich unterscheidet).

Die $n+1$ Gleichungen (75) können nicht unabhängig voneinander sein. Es folgt nämlich für Gleichungen, die von dem Typus der Thieleschen (§ 6) sind, auf bekannte Weise, daß z. B. $F_n^\nu = 0$ erfüllt sein muß, wenn die übrigen n Gleichungen (75), zu denen doch auch die invariante Relation $F_0^\nu = 0$ gehört, erfüllt sind. Es handelt sich dabei eigentlich um folgendes. Für das gegebene Gruppenstück $\mathfrak{M}$ ist $\hat{\gamma}$ der Scharparameter, außerdem unterscheiden sich die in $\mathfrak{M}$ enthaltenen Lösungen voneinander nicht nur in der Epochenkonstante, sondern sie stellen (im Kleinen) verschiedene Bahnkurven dar. Es gibt aber dennoch eine *zwei*parametrige Schar von periodischen Lösungen, der eine Parameter ist $\hat{\gamma}$ (= Jacobische Konstante), der andere die Epoche. Würde man nun annehmen, daß die Gleichungen (75) unabhängig voneinander sind, so würde es sich um $n+1$ Gleichungen für $n+2$ Variable handeln, es würde sich also der Widerspruch ergeben, daß die Epoche nicht willkürlich ist, indem es nur eine *ein*parametrige Schar gibt. — Also muß, wie behauptet, z. B. die Gleichung $F_n^\nu = 0$ eine Folge der n übrigen sein.

Durch die obige Betrachtung ist aber schon angezeigt, wie diese Schwierigkeit zu beheben ist: wir streichen in (75) die überflüssige Gleichung $F_n^\nu = 0$, fügen aber zu den restlichen n Gleichungen eine $(n+1)$-te hinzu, die die Epoche festlegt, und erreichen dadurch, daß zwei Lösungen, die sich voneinander nur in der Epoche unterscheiden, nicht auftreten können. Wir können diese neue Gleichung in der Form $F_*^\nu = 0$ ansetzen, wobei

$$F_*^\nu = F_*^\nu(\nabla\gamma, \nabla\xi_1^0, \ldots, \nabla\xi_n^0)$$

eine in dem Bereiche (72) jedenfalls reguläre Funktion ist. Wir können

z. B., wenn

$$g \equiv \sum_{j=1}^{n} A_j \xi_j + \text{konst.} = 0 \qquad (A_j = \text{konst.})$$

eine feste Gerade des n-dimensionalen reellen ξ-Raumes ist, F_*^V einfach durch $F_*^V \equiv g$ definieren, wenn wir die Epoche dadurch festlegen wollen, daß die Gerade g genau zur Zeit $\tau = 0$ passiert werden soll (bei den Kopenhagener Gruppen kann man entweder die Syzigienachse oder eine einen äquilateralen Librationspunkt enthaltende Gerade als Mannigfaltigkeit g wählen, die freilich nicht immer eine Gerade sein wird). — Es sind nun die $n+1$ Funktionen

$$F_*^V,\ F_0^V,\ F_1^V,\ \ldots,\ F_{n-1}^V,$$

die wir durchlaufend mit

$$f_k = f_k(\vartheta, V\gamma, V\xi_1^0, \ldots, V\xi_n^0) \qquad (k = 0, 1, \ldots, n)$$

bezeichnen wollen, z. B. bei Anwendung auf die Thieleschen Gleichungen (§ 6) gewiß unabhängig voneinander, d. h. keine der $n+1$ Gleichungen $f_k = 0$ folgt identisch aus den n übrigen (insbesondere verschwindet also keine der Funktionen f_k identisch). Wir wollen diese $n+1$ Gleichungen $f_k = 0$ zwischen den $n+2$ Größen $(70)'$ kurz Grundgleichungen nennen; sie liefern nicht nur die notwendigen und hinreichenden Bedingungen der Periodizität und des Erfülltseins der invarianten Relation (6), sie legen vielmehr auch die Epoche fest und sind eben deshalb unabhängig voneinander.

Zieht man also nunmehr nur reelle Werte der Argumente $(70)'$ der Funktionen f_k in Betracht, geht man also von dem komplexen Bereiche (72) zu einem in (72) gelegenen Würfelbereich $\mathfrak{W}$ des reellen $(n+2)$-dimensionalen Raumes über:

$$(80)\quad \mathfrak{W}:\ -\varrho \leqq V\vartheta \leqq \varrho,\ -\varrho \leqq V\gamma \leqq \varrho,\ -\varrho \leqq V\xi_1^0 \leqq \varrho,\ \ldots,\ -\varrho \leqq V\xi_n^0 \leqq \varrho,$$

so repräsentieren unsere Grundgleichungen in dem Bereiche $\mathfrak{W}$ genau $n+1$ reguläre, voneinander verschiedene, genau $(n+1)$-dimensionale Überflächenstücke $\mathfrak{F}_k$ $(k = 0, 1, \ldots, n)$, die wegen (79) den Mittelpunkt O von $\mathfrak{W}$ zum gemeinsamen Punkt haben und einander in O nicht nur berühren, sondern in der Umgebung von O auch weitere gemeinsame Punkte haben, da doch die Punkte (73) voneinander verschieden sind (vgl. S. 343 oben), nach (78) die Grundgleichungen befriedigen und nach (74) gegen O konvergieren. Da alle $n+1$ Überflächenstücke $\mathfrak{F}_k$ regulär, reell, voneinander verschieden und genau von der Dimensionszahl $n+1$ sind, der reelle Bereich $\mathfrak{W}$ aber $(n+2)$-dimensional ist, so müssen also die $n+1$ Überflächenstücke einander in O *durchschneiden*, so daß, wenn ϱ in (80) hinreichend klein gewählt wird, die alle (in $\mathfrak{W}$ liegenden) Lösungen der Grundgleichungen darstellende Schnittfigur der $\mathfrak{F}_k$ aus *endlich vielen, aber*

mindestens zwei von O ausgehenden und einander innerhalb $\mathfrak{W}$ nur in O treffenden, in $\mathfrak{W}$ stetigen, analytischen und mit eventueller Ausnahme des eventuellen algebraischen Verzweigungspunktes O sogar regulär analytischen Kurvenzweigen $\mathfrak{Z}_l$ besteht; vgl. weiter unten den letzten Absatz dieses Paragraphen. — Ist $\mathfrak{Z}_{\bar{l}}$ irgendeiner dieser Zweige, so gibt es nach bekannten, auf Puiseux zurückgehenden Sätzen mindestens einen, *von* $\mathfrak{Z}_{\bar{l}}$ *verschiedenen* Zweig $\mathfrak{Z}_{\bar{l}^*}$, der eine unmittelbare analytische Fortsetzung von $\mathfrak{Z}_{\bar{l}}$ ist. Die Möglichkeit für die analytische Fortsetzung des gegebenen Gruppenstückes (31), (32) für $\lambda > \bar{\bar{\lambda}}$, d. h. über $\bar{\bar{\lambda}}$ hinaus, wird also bewiesen sein, wenn wir die Existenz einer in (32) gelegenen Zahl $\bar{\lambda}'$ beweisen können, die derart ist, daß das dem Teilintervalle

$$\bar{\lambda}' < \lambda < \bar{\bar{\lambda}} \tag{81}$$

von (32) entsprechende Stück $\mathfrak{M}'$ des zu (32) gehörigen Gruppenstückes $\mathfrak{M}$ durch die Punkte genau eines Zweiges $\mathfrak{Z}_{\bar{l}}$ der Schnittfigur $\Sigma\mathfrak{Z}_l$ eineindeutig repräsentiert wird.

Es bezeichne $P(\lambda_\nu)$ den zu λ_ν gehörigen Punkt (73) des $(n+2)$-dimensionalen Raumes der reellen Variablen (70)′ und nehmen wir die Bezeichnung $P(\lambda)$ überhaupt bei allen in dem Kontinuum (32) liegenden λ-Werten, d. h. nicht nur für die Folge $\lambda = \lambda_\nu$, in Anspruch, so daß also der Punkt $P(\lambda)$ auf der Schnittfigur $\Sigma\mathfrak{Z}_l$ der $\mathfrak{F}_k$ liegen muß, sobald er in dem Würfel (80) liegt. Wegen (74) liegt $P(\lambda_\nu)$ bei hinreichend großem ν gewiß in $\mathfrak{W}$, also auf $\Sigma\mathfrak{Z}_l$; es bezeichne $\mathfrak{Z}_{l_\nu}$ denjenigen Zweig $\mathfrak{Z}_l$, der den Punkt $P(\lambda_\nu)$ enthält. Da es nur endlich viele $\mathfrak{Z}_l$, aber unendlich viele $P(\lambda_\nu)$ gibt, so gibt es mindestens einen Zweig $\mathfrak{Z}_{\bar{l}}$, der *unendlich viele* Punkte $P(\lambda_\nu)$ enthält. Offenbar können wir voraussetzen, daß *alle* Punkte $P(\lambda_\nu)$ auf diesem festen Zweige $\mathfrak{Z}_{\bar{l}}$ liegen: um dies zu erreichen, braucht man nur aus der Folge $\{\nu\}$ der natürlichen Zahlen eine passende Teilfolge $\{\nu'\}$ auszusondern und dann den Zeiger ν' der Auswahlfolge $\{\lambda_{\nu'}\}$ einfach durch ν zu bezeichnen. (52) gilt offenbar auch für die derart festgelegte neue Folge $\{\lambda_\nu\}$:

$$P(\lambda_\nu) \to O, \tag{82}$$

$$\lambda < \lambda_\nu < \lambda_{\nu+1} \to \bar{\bar{\lambda}}, \tag{83}$$

außerdem liegen alle $P(\lambda_\nu)$ auf einem und demselben Zweige:

$$\mathfrak{Z}_{l_\nu} = \mathfrak{Z}_{\bar{l}} \qquad (\nu = 0, 1, \ldots). \tag{84}$$

Bewegt sich nun λ in dem Intervall $\lambda_\nu \leqq \lambda \leqq \lambda_{\nu+1}$, so befindet sich $P(\lambda)$ für $\lambda = \lambda_\nu$ und für $\lambda = \lambda_{\nu+1}$ wegen (84) auf einem und demselben Zweige $\mathfrak{Z}_{\bar{l}}$. Würde also $P(\lambda)$ zwischen $\lambda = \lambda_\nu$ und $\lambda = \lambda_{\nu+1}$ den Zweig $\mathfrak{Z}_{\bar{l}}$ verlassen [was dann nur entweder über O hindurch oder außerhalb des Würfels $\mathfrak{W}$ geschehen könnte], oder würde sich $P(\lambda)$ von $P(\lambda_\nu)$ bis $P(\lambda_{\nu+1})$

nicht monoton, d. h. nicht ohne Richtungsändetung auf dem zwischen $P(\lambda_\nu)$ und $P(\lambda_{\nu+1})$ gelegenen Stücke von $\mathfrak{Z}_{\bar{\lambda}}$ bewegen, so würde es auf $\mathfrak{Z}_{\bar{\lambda}}$ mit Rücksicht auf die Stetigkeit und den offenbar abteilungsweise monotonen Charakter dieser Bewegung wegen (82), (83) gewiß ein Stück geben, das von $P(\lambda)$ auf eine „gespiegelte" Weise (vgl. S. 336), also nach einer Umkehr in einem Punkte von $\mathfrak{Z}_{\bar{\lambda}}$ zweimal nacheinander beschrieben wird, was aber der am Schluß von § 3 kursiv gesetzten Festsetzung widerspricht. — Damit ist die Existenz eines Intervalles (81) von dem im vorigen Absatz beschriebenen Charakter bewiesen: wir haben nur $\bar{\lambda}' = \lambda_0$ zu setzen.

Insbesondere sind also die Funktionen (31) in dem offenen Intervall (32) nicht nur stetig, sondern auch gleichmäßig stetig. Daß die Periode $\hat{T}(\lambda)$ für $\bar{\lambda} < \lambda < \bar{\bar{\lambda}}$ von Null verschieden ist, versteht sich von selbst, da sie in diesem *offenen* Intervalle von vornherein als die Periode einer *gegebenen* periodischen Lösung definiert war. Es wäre aber denkbar, daß die zu $\lambda = \bar{\bar{\lambda}}$ gehörige Bahn nicht mehr periodisch ist, indem $\hat{T}(\bar{\bar{\lambda}})$, d. i. die unter (51), (52) mit Θ^* bezeichnete Zahl, verschwindet. Doch handelt es sich dabei nur um eine scheinbare Schwierigkeit, da in dem n-dimensionalen Raume der reellen Koordinaten ξ_i die zu $\lambda (< \bar{\bar{\lambda}})$ gehörige Bahnkurve geschlossen ist und für $\lambda \to \bar{\bar{\lambda}} - 0$ mit Rücksicht auf den Cauchy-Poincaréschen Stetigkeitssatz (vgl. § 1) auf jedem endlichen τ-Intervall gegen die zu $\lambda = \bar{\bar{\lambda}}$ gehörige Bahnkurve konvergiert; da andererseits die Periode $\hat{T}(\lambda)$ bei Vornahme des Grenzüberganges $\lambda \to \bar{\bar{\lambda}} - 0$ beschränkt bleibt, so muß auch die zu $\lambda = \bar{\bar{\lambda}}$ gehörige Bahnkurve geschlossen, also die Lösung periodisch sein. Auf die Problematik des Poincaréschen zweiten Genres (Doppelbahnen, kommensurable charakteristische Exponenten) soll erst in einer Fortsetzung dieser Arbeit eingegangen werden. — Soviel wäre übrigens von vornherein klar gewesen, daß die Funktion $\hat{T}(\lambda)$ für $\lambda > \bar{\bar{\lambda}}$ bei hinreichend kleinem $\lambda - \bar{\bar{\lambda}} > 0$ nicht verschwinden kann, da sie sonst identisch verschwinden würde, während ihre (auf $\hat{\gamma}$ bezogene) direkte analytische Fortsetzung für $\lambda < \bar{\bar{\lambda}}$ nicht verschwindet.

Es sei hier noch bemerkt, daß man an Stelle von O einen beliebigen, in der Umgebung von O liegenden Punkt $P(\lambda)$ zugrunde legen könnte, da ε und δ_ε unabhängig von λ sind.

Wir haben bewiesen, daß, wenn $m(\lambda)$ [in (46)] bei Vornahme des Grenzüberganges $\lambda \to \bar{\bar{\lambda}} - 0$ nicht gegen $+\infty$ konvergiert [wobei also der a priori denkbare Fall einer „wesentlichen Singularität", scheinbar eben die Hauptschwierigkeit des Problems, mitberücksichtigt ist], das Gruppenstück $\mathfrak{M}$ über $\bar{\bar{\lambda}}$ hinaus eine reelle analytische Fortsetzung zuläßt. Nehmen wir nun an, daß $m(\lambda)$ beschränkt bleibt, wie weit man auch für $\lambda > \bar{\bar{\lambda}}$ die analytische Fortsetzung treiben soll. Ist dann $\bar{\bar{\lambda}} \leqq \lambda < \Lambda$ oder $\bar{\bar{\lambda}} < \lambda \leqq \Lambda$ das Parametergebiet des *ganzen*, auf $\bar{\bar{\lambda}}$ folgenden Gruppenstückes, wobei

mit Rücksicht auf die Willkür der Parameterwahl offenbar $\Lambda < +\infty$ angenommen werden kann, so kann die Gruppe nach dem Vorangehenden auch über Λ hinaus fortgesetzt werden. Dies steht aber im Widerspruch mit der Definition von Λ, wonach sich bei Vornahme des Grenzüberganges $\lambda \to \Lambda - 0$ das *ganze*, nach $\bar{\bar{\lambda}}$ folgende Gruppenstück ergeben soll. Der Widerspruch löst sich dann und nur dann, wenn man beachtet, daß der Begriff des *ganzen*, nach $\bar{\bar{\lambda}}$ folgenden Gruppenstückes bei in sich geschlossenen Bahnmannigfaltigkeiten und nur bei diesen illusorisch ist. Ich nenne dabei die Gruppe dann und nur dann „in sich geschlossen", wenn das $(n+1)$-tupel der analytischen Funktionen (44) des reellen Arguments $\hat{\gamma}$, auf den in Betracht kommenden Streckenkomplexen der reellen Gebiete ihrer Riemannschen Flächen als Funktionen von $\hat{\gamma}$ aufgefaßt, periodisch in $\hat{\gamma}$ ist. — Es ist freilich denkbar, daß auf der Riemannschen Fläche ein geschlossener Weg beschrieben wird, ohne daß dabei der Ausdruck (46) beschränkt bliebe, wobei dann die Gruppe in dem vorigen Sinne nicht notwendig geschlossen ist.

Es gilt also die folgende Alternative: Jede Gruppe periodischer Lösungen ist entweder in sich geschlossen, oder aber sie klingt auf eine „natürliche" Weise, nämlich dadurch ab, daß entweder die Periode, oder die „Energie" $\hat{\gamma}$, oder die Bahndimensionen nicht beschränkt bleiben, oder endlich eine unbegrenzte Annäherung an eine der (fünf) Gleichgewichtslösungen stattfindet.

Der oben (S. 344) herangezogene Satz, wonach es höchstens endlich viele, aber mindestens zwei Zweige $\mathfrak{Z}_l$ geben muß (S. 345), folgt am einfachsten durch den Schluß von p auf $p+1$ der p-dimensionalen Geometrie. Einen an die Puiseuxschen Sätze anknüpfenden (durchaus korrekten) funktionentheoretischen Beweis gab Poincaré an der oben wegen anderen Dingen beanstandeten Stelle der Méth. Nouv.

§ 6.

Anwendung auf die Thieleschen Bewegungsgleichungen.

Man setze in (2) die Zahl $n = 4$ und definiere die $4+1$ ganzen Funktionen (1), (5) wie folgt, wobei μ eine numerische Konstante bedeutet:

$$G_1 = \xi_3,$$

$$G_2 = \xi_4,$$

$$G_3 = (\mathfrak{Cof}\, 2\xi_2 - \cos 2\xi_1)\,\xi_4 + \frac{1}{4}\sin 4\xi_1 - \frac{1}{2}\gamma \sin 2\xi_1$$
$$+ (1-2\mu)\left(\frac{1}{4}\sin \xi_1\, \mathfrak{Cof}\, 3\xi_2 - \frac{3}{4}\sin 3\xi_1\, \mathfrak{Cof}\, \xi_2 - 8\sin \xi_1\right),$$

$$G_4 = (-\mathfrak{Cof}\,2\xi_2 + \cos 2\xi_1)\,\xi_3 + \frac{1}{4}\,\mathfrak{Sin}\,4\xi_2 - \frac{1}{2}\,\gamma\,\mathfrak{Sin}\,2\xi_2$$
$$+ (1-2\mu)\left(\frac{1}{4}\,\mathfrak{Sin}\,\xi_2\cos 3\xi_1 - \frac{3}{4}\,\mathfrak{Sin}\,3\xi_2\cos\xi_1\right) + 8\,\mathfrak{Sin}\,\xi_2,$$

$$\Gamma = \xi_3^2 + \xi_4^2 - \frac{1}{8}\,(\mathfrak{Cof}\,4\xi_2 - \cos\xi_1) + \frac{1}{2}\,\gamma\,(\mathfrak{Cof}\,2\xi_2 - \cos 2\xi_1)$$
$$+ (1-2\mu)\left(\frac{1}{2}\cos\xi_1\,\mathfrak{Cof}\,3\xi_2 - \frac{1}{2}\cos 3\xi_1\,\mathfrak{Cof}\,\xi_2 - 16\cos\xi_1\right).$$

Dann ist (2) das System der Bewegungsgleichungen des restringierten Dreikörperproblems in den Thieleschen Regularisierungsvariablen, und (6) ist das Jacobische Integral[16]). Die in den numerischen Untersuchungen der Kopenhagener Sternwarte durch

$$E, \quad F, \quad \frac{dE}{d\psi}, \quad \frac{dF}{d\psi}, \quad \psi, \quad K$$

bezeichneten Größen sind dabei bez. mit

$$\xi_1, \quad \xi_2, \quad \xi_3, \quad \xi_4, \quad \tau, \quad \gamma$$

zu identifizieren; μ ist der Massenprozentsatz, dem ein fester Wert erteilt wird (in fast allen Kopenhagener Rechnungen ist $1-2\mu=0$). Unter $T = T_K$ die auf ψ bezogene Periode einer periodischen Bahn $E = E(\psi)$, $F = F(\psi)$ verstanden, führe man das Quadrat der maximalen Entfernung

$$R_K = \max_{0 \leqq \psi \leqq T_K} \{[E(\psi)]^2 + [F(\psi)]^2\}$$

der Bahn von dem Ursprung der Koordinatenebene ein. Da das Energieintegral (6) die Form

$$\left(\frac{dE}{d\psi}\right)^2 + \left(\frac{dF}{d\psi}\right)^2 = f(K; E, F)$$

hat, wobei $f(\gamma, \xi_1, \xi_2)$ die ganze Funktion $\Gamma(\gamma; \xi_1, \xi_2, \xi_3, \xi_4) - \xi_3^2 - \xi_4^2$ bedeutet, so ist die unter (45) durch $r(\lambda)$ bezeichnete Schranke gewiß nicht größer als $R_K + \bar{f}(R_K, R_K, |K|)$, wenn $\bar{f}$ die beste Majorante der Potenzreihe f, also ebenfalls eine ganze Funktion von drei Veränderlichen bezeichnet, so daß die unter (46) durch $m(\lambda)$ bezeichnete Schranke nicht größer ist als

$$R_K + \bar{f}(R_K, R_K, |K|) + |K| + T_K,$$

sie kann daher, als Funktion des Scharparameters $K = \gamma$ aufgefaßt, nur dann gegen $+\infty$ konvergieren, wenn

$$R_K + |K| + T_K$$

---

[16]) T. N. Thiele, Astr. Nachr. **138** (1895), S. 1 ff. [für $1-2\mu=0$]; C. Burrau, Vierteljahresschr. der Astr. Ges. **41** (1906), S. 261 ff. [für beliebiges μ]. — Wegen Literatur s. Math. Zeitschr. **32** (1930), S. 675. Vgl. ferner J. Fischer-Petersen, Astr. Nachr. **200** (1915), S. 385 ff. [= Publ. Observ. Nr. 22]. Wegen einer Art Regularisierung in dem unendlich fernen Punkte der kartesischen Bahnebene s. B. O. Koopman, a. a. O. XVII.

das tut, wenn also entweder die Bahndimensionen oder die Energie oder endlich die Periode nicht beschränkt bleiben.

Jede Gruppe von periodischen Lösungen des restringierten Dreikörperproblems ist also entweder in sich geschlossen oder sie erleidet entweder in einer der fünf Gleichgewichtslösungen oder durch Unendlichwerden entweder der Bahndimensionen oder der Periode oder der Jacobischen Konstante ein „natürliches" Absterben. In dieser Alternative ist von Stoßbahnen oder Nichtstoßbahnen gar nicht die Rede, die Gruppe kann also über beliebig viele Stoßbahnen hindurch ohne weiteres fortgesetzt werden. Würde man Stoßbahnen nicht zulassen, so wäre der Satz über den natürlichen Abschluß bzw. die Geschlossenheit ganz falsch; es war auch ein leitendes Prinzip der Strömgrenschen numerischen Rechnungen, bei einer periodischen Stoßbahn nicht haltzumachen; *mathematisch* war aber bisher u. a. die Möglichkeit einer Fortsetzung über eine — in dem Thieleschen regularisierenden Variablensystem ganz harmlose, d. h. den Nichtstoßbahnen gegenüber auf keine Weise ausgezeichnete — Stoßbahn der Gruppe hindurch nicht bewiesen. —

In dem Hillschen parallaxenlosen Grenzfall, wobei es nur eine einzige Massensingularität gibt, kommt man an Stelle der Thieleschen Transformation mit der Levi-Civitaschen[17]) durch; unsere Ergebnisse gelten also auch in dem Hillschen Grenzfall.

Wir haben oben γ mit der Jacobischen Konstante identifiziert und den Massenprozentsatz fest gewählt, um es mit *Gruppen* zu tun zu haben. Ebensogut könnte man freilich γ mit dem Massenprozentsatz identifizieren und der Jacobischen Konstante einen festen Wert erteilen. Die sich so ergebenden Sätze sind, wie in einer Fortsetzung dieser Arbeit gezeigt wird, einerseits für die Poincaré-Birkhoffsche Problematik der Sorten, andererseits für das Problem der nichtinfinitesimalen Librationen von Bedeutung.

The Johns Hopkins University, December 11th, 1930.

---

[17]) Gemeint ist

$$\xi = E^2 - F^2, \qquad \eta = 2EF, \qquad dt = 4(E^2 + F^2)\,d\psi;$$

vgl. z. B. XVI, S. 275 und S. 314.

(Eingegangen am 7. Januar 1931.)

Grundlagen einer Genealogie der periodischen Bahnen im restringierten Dreikörperproblem.

Zweite Mitteilung.

Sortengenealogie, Hekubakomplex und Gruppenfortsetzung.

Von

Aurel Wintner in Baltimore, Md.

Inhalt.

		Seite
Einleitung		350
§ 1.	Reduktionsformeln der kanonischen, der Darwinschen und der Kopenhagener Jacobischen Konstanten	353
§ 2.	Der a priori zulässige Phasenraum	359
§ 3.	Nähere Bestimmung der a priori zulässigen Gebiete	367
§ 4.	Die Sortenmannigfaltigkeit	373
§ 5.	Die Massenschranke	377
§ 6.	Das Prinzip der Massenfixierung	385
§ 7.	Die Gruppe g und der Hekubakomplex	389
§ 8.	Bestätigung der Theorie durch Reduktion der Darwinschen Satellitenfamilien auf die Kopenhagener Gruppe g	
§ 9.	Die übrigen Gruppen mit Sortenursprung	396
§10.	Die isoperiodischen Sorten	400

Einleitung.

Die vorliegende Arbeit bringt eine wesentliche Verschärfung und feinere Ausgestaltung gewisser in einer vorläufigen Mitteilung[1]) in Angriff genommenen Untersuchungen, die zur Behandlung der in der ersten Mitteilung[2]) nur berührten Probleme in einer späteren Mitteilung notwendig sein werden, und verfolgt, sich auf die erste Mitteilung stützend, einen doppelten Zweck.

---

[1]) Sitzungsber. der Sächs. Akad. d. Wiss. zu Leipzig **82** (1930), S. 2–56.

[2]) Math. Zeitschr. **34** (1931), S. 321–349 [= Publ. Købhn. Observ. Nr. 75].

Einerseits wird zunächst gezeigt, daß die Poincarésche Theorie[3][4][5] der periodischen Sorten des restringierten Dreikörperproblems[6] mathematisch nicht in allen ihren Teilen haltbar ist. — So sind z. B. die Poincaréschen isoperiodischen zweiten Sorten eben für hinreichend kleine Werte der Eigenexzentrizität nicht vorhanden; der Existenzbeweis von Poincaré beruht, ebenso wie eine analoge Untersuchung von Schwarzschild[7], auf einem Mißverständnis. Ferner ist in der Poincaréschen Sortengenealogie[5] manches nicht ganz streng begründet, nicht selten läßt sich auch das Schlußresultat selbst nicht mehr aufrechterhalten. — Ich operiere dabei mit gewissen Diagrammen in der Mannigfaltigkeit aller störungstheoretisch, d. h. im Rahmen der Poincaréschen Sortenproblematik überhaupt erreichbaren periodischen Lösungen aller, d. h. zu beliebigen Werten des Massenprozentsatzes μ gehörigen restringierten Dreikörperprobleme. In diesen Phasenraum werden auch einige Ergebnisse der Rendiconti-Arbeit von Herrn Birkhoff[8] eingebettet. Die Diagramme, die „Massenschranken", fixieren die obere Grenze der störungstheoretisch, d. h. im Rahmen der Poincaréschen Fragestellungen überhaupt in Betracht kommenden Werte von μ als eine Funktion des Ortes in dem Phasenraume der ungestörten periodischen Lösungen, ermöglichen erst eine systematische Sortengenealogie, gewähren ferner wesentliche Einsicht in den mathematisch zur Zeit durchaus problematischen Lückenkomplex kritischer (d. h. synodisch ganzzahliger) Kommensurabilitäten, indem die Massenschranken in Evidenz setzen, daß es sich — bei beliebiger (also noch so kleiner), aber *fester* Wahl der störenden Jupitermasse μ — in einer *vollen Umgebung* der synodisch exakt ganzzahligen Kommensurabilitätsstellen (die in dem Texte dementsprechend Lückenzentren heißen) zunächst gar nicht um die Stabilität[9], sondern einfach um die Existenz der periodischen Sortenlösungen handelt.

---

[3]) H. Poincaré, Méth. nouv. de la méc. cél. **1** (1892).

[4]) H. Poincaré, Méth. nouv. de la méc. cél. **3** (1899).

[5]) H. Poincaré, Bull. Astr. **19** (1902), S. 177—198.

[6]) Mit dem sogenannten reduzierten Dreikörperproblem, wobei die Jupiterexzentrizität nicht verschwindet und es daher kein Jacobisches Integral gibt, beschäftige ich mich hier *nicht*. Über dieses Problem liegen die bekannten Untersuchungen von Hill [Coll. Math. Works **4** (1907), S. 169 ff.] und Wilkens [Ergänzungshefte zu den Astr. Nachr. 1905; Astr. Nachr. **195** (1913), S. 385 ff. usw.], ferner Stabilitätsrechnungen von Klose [Astr. Nachr. **218** (1923), S. 401 ff.] sowie gewisse Ansätze von W. W. Heinrich vor; im übrigen möge man den Enzyklopädieartikel von Samter vergleichen.

[7]) K. Schwarzschild, Astr. Nachr. **147** (1896), S. 17—24.

[8]) G. D. Birkhoff, Rend. Palermo **39** (1915), S. 265—334.

[9]) Wegen der Literatur des Stabilitätsproblems vgl. a. a. O. [1]), S. 51 ff.

Andererseits wird untersucht, was denn die Poincarésche Sortenproblematik (und damit jede Anschmiegungs-, d. h. Störungstheorie der periodischen Lösungen) überhaupt leisten *kann*. Das Ergebnis dieser Untersuchung, die einerseits an Hand des Lückenkomplexes und andererseits auf Grund der Kopenhagener Arbeiten[10]),[11]) bzw. der ersten Mitteilung geführt wird, besteht darin, daß — genealogisch betrachtet — der Sortentheorie eine entscheidende Bedeutung nicht zugeschrieben werden kann. Denn es handelt sich dabei um Gedankengänge, die den eigentlichen Problemen nicht gerecht werden *können*, und zwar bereits deshalb nicht, weil z. B. die Theorie der Sorten auch in dem besten Fall nur ein kleines (im Sinne der ersten Mitteilung „dynamisch selbstverständliches") Stück einer Gruppe liefern kann. Diese *entwickelt* sich aber nach Passieren der Massenschranke der Sortentheorie *oberhalb* dieser oberen Massenschranke und über die kritischen Lückengebiete hindurch *selbständig weiter* so lange, als es in der ersten Mitteilung[2]) präzisiert worden ist (Strömgrensches Abschlußprinzip[10])). Anders ausgedrückt, die Probleme im großen nehmen erst in einem Phasengebiet ihren Anfang, das für die Sortentheorie einfach nicht existiert, das nämlich mittels sortenmäßiger analytischer Fortsetzung *nicht erreichbar* ist, indem es überhaupt nicht in dem Phasenraume der Sortenproblematik liegt. Dem Numeriker sind diese Verhältnisse in einem anderen Gewand seit langem bekannt[12]), [12a]), sie traten dadurch zum Vorschein, daß der analytische Apparat der Sortenproblematik praktisch versagte. Nun liegt aber bei dieser Erscheinung nicht etwa eine nur rechentechnische Schwierigkeit vor. Vielmehr handelt es sich um eine rein funktionentheoretische Unmöglichkeit einer analytischen Fortsetzung im Rahmen der Sortentheorie, während die analytische Fortsetzung an sich ohne weiteres möglich ist, aber aus dem Phasenraume der Sortentheorie zwangsweise hinausführt. Die direkte Hillsche Behandlung ist also, wo es zur Zeit überhaupt möglich ist — und es gibt solche Fälle —, der Anschmiegungsmethode gegenüber bereits aus diesem Grunde unbedingt zu bevorzugen, ganz abgesehen davon, daß sie zugleich die Methode der Praxis ist und dementsprechend auch explizite Konvergenzgebiete zu berechnen gestattet. Doch kann auf diese nur scheinbar selbstverständlichen

---

[10]) Vgl. die auf S. 322 ff. der ersten Mitteilung zusammengestellten zusammenfassenden Berichte von E. Strömgren.

[11]) Zur Orientierung vgl. man z. B. E. Strömgren, Ergebnisse der exakten Naturwissenschaften 4 (1925), S. 233 ff., oder H. Samter, Enzykl. d. Math. Wiss. VI 2_1, S. 974 ff.

[12]) E. Strömgren, Tre Aartier Celest Mekanik paa Københavns Observatorium, Kopenhagen, 1923, S. 11.

[12a]) Vgl. ferner die Ausführungen von E. Strömgren in der Vierteljahrsschrift der Astr. Gesellsch. 1930 [= Publ. Købn. Nr. 72, S. 4–5].

und — trotz der Grundschwierigkeit der älteren Mondtheorien, die eben Störungstheorien waren — nicht einmal in der allerletzten Zeit beachteten Dinge erst weiter unten eingegangen werden. — Im übrigen handelt es sich um die Strömgrenschen Gruppen[11]) f, g, h, i, l, m (bei beliebigem μ).

§ 1.

Reduktionsformeln der kanonischen, der Darwinschen und der Kopenhagener Jacobischen Konstanten.

Die „kanonischen" Einheiten in dem restringierten Dreikörperproblem sind wie folgt festgelegt.

Die Masseneinheit ist die Gesamtmasse, d. h. die Massen der beiden, den restringierten Aufpunkt anziehenden, von diesem aber nicht gestörten Massenpunkte können gleich μ und $1-\mu$ gesetzt werden, wenn μ den „Massenprozentsatz" bedeutet, der an sich seinem Komplement $1-\mu$ gegenüber auf keine Weise ausgezeichnet ist und nur der Bedingung

$$0 \leqq \mu \leqq 1 \tag{1}$$

zu genügen hat, wobei also $\mu = 0$ und $\mu = 1$ auch erlaubt sind. Diese beiden Massenpunkte wirken aufeinander nach dem Newtonschen Gesetz (wobei der Grenzfall $\mu = 0$ bzw. $\mu = 1$ auf die übliche Weise zu definieren ist), und die Integrationskonstanten dieser Bewegung sind, im Gegensatz zu dem sogenannten reduzierten Dreikörperproblem, in dem restringierten Dreikörperproblem derart gewählt, daß μ und $1-\mu$ (in einem inertialen Achsenkreuz) um ihren Schwerpunkt O Keplersche *Kreise* beschreiben, so daß also die Verbindungsgerade dieser beiden Massenpunkte, d. h. die Syzygienachse, mit konstanter Drehgeschwindigkeit in einer zeitunabhängigen inertialen Ebene rotiert. Der Umlaufssinn in dieser Ebene wird dadurch festgelegt, daß die Drehgeschwindigkeit der Syzygienachse positiv, und die Zeiteinheit dadurch, daß diese Drehgeschwindigkeit gleich 1 ist. Als Längeneinheit wählt man endlich die zeitunabhängige Entfernung der beiden Massen μ, $1-\mu$. Mit Rücksicht auf das dritte Keplersche Gesetz ist dann auch die Gravitationskonstante gleich 1.

Die Anfangslage und die Anfangsgeschwindigkeit des dritten Körpers, d. h. des restringierten Aufpunktes werden in der gemeinsamen Bahnebene von μ und $1-\mu$ gewählt, so daß also die ganze Bewegung aller drei Körper in einer zeitunabhängigen Ebene erfolgt. Als Anfangspunkt dieser inertialen Ebene können wir den Schwerpunkt O wählen und die Richtung der x-Achse dadurch festlegen, daß zur Zeit $t = 0$ die Ordinate der Masse $1-\mu$ (also auch diejenige von μ) gleich Null, ihre Abszisse aber $\geqq 0$ (also diejenige von μ offenbar $\leqq 0$) ist. Da $x = 0$, $y = 0$ der Schwerpunkt ist und die Entfernung der beiden Massen μ, $1-\mu$ die

Einheit beträgt, so sind dann die Koordinaten von $1-\mu$ und μ zur Zeit $t=0$ offenbar $x=\mu$, $y=0$ und $x=-(1-\mu)$, $y=0$. Dieses Koordinatensystem (x, y) nennen wir siderisches Achsenkreuz. Von diesem gehen wir mittels der orthogonalen Substitution

$$x=\xi\cos t-\eta\sin t, \qquad y=\xi\sin t+\eta\cos t \tag{2}$$

zu einem Koordinatensystem (ξ, η) über, das synodisches Achsenkreuz heißt, um den Schwerpunkt ($x=y=0$, d. h. $\xi=\eta=0$) mit der konstanten Drehgeschwindigkeit ($=1$) der Syzygienachse rotiert und zur Zeit $t=0$ mit dem siderischen Achsenkreuz zusammenfällt. Die ξ-Achse ist also nicht nur für $t=0$, sondern ständig die Syzygienachse, und die beiden Massenpunkte $1-\mu$, μ haben in dem synodischen Achsenkreuz die zeitunabhängigen Koordinaten

$$\begin{aligned} 1-\mu&: \quad \xi=\mu, \quad \eta=0; \\ \mu&: \quad \xi=-(1-\mu), \quad \eta=0. \end{aligned} \tag{3}$$

Die Differentialgleichungen für die Koordinaten ξ, η des dritten Körpers sind daher

$$\ddot{\xi}-2\dot{\eta}=\frac{\partial\Omega}{\partial\xi}, \qquad \ddot{\eta}+2\dot{\xi}=\frac{\partial\Omega}{\partial\eta} \qquad \left(\dot{} =\frac{d}{dt}\right) \tag{4}$$

mit

$$2\Omega\equiv 2\Omega_\mu(\xi,\eta)=\xi^2+\eta^2+\frac{2\mu}{|\sqrt{(\xi+1-\mu)^2+\eta^2}|}+\frac{2(1-\mu)}{|\sqrt{(\xi-\mu)^2+\eta^2}|}, \tag{5}$$

indem Ω die Kräftefunktion der konservativen, d. h. Zentrifugal- und Gravitationskräfte (d. h. aller Kräfte mit Ausschluß der in (4) rechter Hand angeschriebenen Coriolisterme) ist und die beiden Nenner in (5) die auf die Zentra (3) bezogenen bipolaren Koordinaten bedeuten. Aus (4) folgt das Jacobische Integral

$$\dot{\xi}^2+\dot{\eta}^2=2\Omega-C, \tag{6}$$

wobei

$$C=C_\mu\equiv C_{1-\mu} \tag{7}$$

eine Integrationskonstante bedeutet. Die notwendige und hinreichende Bedingung dafür, daß bei gegebenem Massenprozentsatz μ und bei gegebener Jacobischer Konstante (7) bei willkürlicher Wahl der drei übrigen Integrationskonstanten die (synodische) Geschwindigkeit für irgendeinen Zeitpunkt verschwindet, besteht nach (6), (5) darin, daß die zu diesem Zeitpunkt gehörigen Koordinaten ξ, η die Gleichung

$$2\Omega_\mu=C \qquad (C=C_\mu) \tag{8}$$

befriedigen. (8) ist mit (6) dann und nur dann für alle t identisch, wenn es sich um eine Gleichgewichtslösung handelt, d. h. wenn die Anfangs-

werte ξ_0, η_0 der Koordinaten die beiden Gleichungen

$$\frac{\partial}{\partial \xi}\Omega_\mu(\xi_0, \eta_0) = 0, \qquad \frac{\partial}{\partial \eta}\Omega_\mu(\xi_0, \eta_0) = 0 \tag{9}$$

befriedigen und die Anfangswerte der Geschwindigkeiten gleich Null sind. Für eine solche Gleichgewichtslösung ist die Jacobische Konstante nach (6), (5) gleich

$$2\,\Omega_\mu(\xi_0, \eta_0). \tag{10}$$

Ist $0 < \mu < 1$ [für $\mu = 0$ und $\mu = 1$ liegt die Sache anders; vgl. S. 360 weiter unten], so hat das Gleichungssystem (9) bekanntlich genau fünf (von μ abhängige) Lösungen ξ_0, η_0; drei dieser Librationspunkte liegen auf der Syzygienachse $\eta = 0$ und sind voneinander durch die beiden Punkte (3) getrennt (kollineare Librationspunkte), die beiden übrigen bilden mit den beiden Punkten (3) je ein gleichseitiges Dreieck, haben also die Koordinaten

$$\xi_0 = \mu - \frac{1}{2}, \qquad \eta_0 = \pm \frac{\sqrt{3}}{2} \tag{11}$$

(äquilaterale Librationspunkte). Die Jacobische Konstante der äquilateralen Gleichgewichtslösungen ist nach (11), (10) und (5) gleich

$$3 + \mu(\mu - 1), \tag{12}$$

also z. B. für $\mu = \frac{1}{11}$ oder für $\mu = \frac{10}{11}$ gleich

$$3 - \frac{10}{121} \tag{12a}$$

gleichund für $\mu = \frac{1}{2}$

$$\frac{11}{4}. \tag{12b}$$

Alle fünf Librationspunkte sind mit Rücksicht auf die Definition (9) dadurch charakterisiert, daß die über die (ξ, η)-Ebene ausgebreitete Fläche $\zeta = 2\,\Omega_\mu(\xi, \eta)$ in diesen eine zu $\zeta = 0$ parallele Tangentenebene hat. Die Funktion ζ von ξ und η hat nach einem Satze von Herrn Plummer[13]) (bei jedem von 0 und 1 verschiedenen, fest gewählten Werte von μ) in allen drei kollinearen Librationspunkten einen Sattelpunkt. Ein Maximum oder ein Minimum von ζ ist also nur in den beiden äquilateralen Librationspunkten möglich, in welchen aber ζ einen und denselben Wert (12) annimmt. Folglich ist (12) das absolute Minimum von ζ, da die Funktion ζ jedenfalls ein Minimum haben muß, indem sie nach (5) durchweg > 0 ist und nicht nur in den beiden Massen (3), sondern auch für $\xi^2 + \eta^2 \to +\infty$ unendlich wird. Schneidet man die Fläche ζ mit der Ebene $\zeta = C_\mu$, so ist die Nullgeschwindigkeitskurve (8), sofern sie reell ist, die Projektion

[13]) H. C. Plummer, Monthly Notices of the Royal Astr. Soc. 62 (1902), S. 16–17.

der Schnittfigur auf die Ebene $\zeta = 0$. Daraus folgt einerseits, daß die algebraische Kurve (8) nur so lange auch reelle Punkte hat, als C_μ nicht kleiner ist als das Minimum (12) von ζ, und andererseits folgt leicht, daß ihre reellen Zweige durchweg *geschlossen* sein müssen [in dem Hillschen parallaxenlosen Grenzfall des restringierten Dreikörperproblems[14]), wobei ζ, unabhängig von μ, gleich

$$3\xi^2 + \frac{2}{|\sqrt{\xi^2+\eta^2}|} \tag{13}$$

ist, ist dies nicht der Fall, (13) konvergiert aber auch, im Gegensatz zu (5), für $\xi^2+\eta^2 \to +\infty$ keinesfalls gegen $+\infty$; übrigens hat Hill dieselben Einheiten wie wir]. Wir notieren noch, daß $2\,\Omega_\mu(0,0)$ nach (5) gleich

$$2\left(\frac{\mu}{1-\mu} + \frac{1-\mu}{\mu}\right), \tag{14}$$

also insbesondere für $\mu = \frac{1}{2}$ gleich

$$4 \tag{14'}$$

ist, und daß andererseits $\xi = 0$, $\eta = 0$ im Falle $\mu = \frac{1}{2}$ den zwischen den beiden Massen liegenden kollinearen Librationspunkt darstellt, daß also (14)′ mit Rücksicht auf (10) die Jacobische Konstante der zu diesem Librationspunkt und zu $\mu = \frac{1}{2}$ gehörigen Gleichgewichtslösung sein muß.

Wir wollen nun das durch (6) definierte C mit den sonstigen in der Literatur üblichen Bestimmungen der Jacobischen Konstante vergleichen.

Herr Birkhoff[15]) benutzt die obigen Einheiten und Koordinaten, seine Jacobische Konstante, die wir mit $B = B_\mu$ bezeichnen wollen (bei Herrn Birkhoff heißt sie C), ist aber dennoch nicht unser $C = C_\mu$, es ist vielmehr

$$B_\mu = C_\mu + \mu(1-\mu) \tag{15}$$

(so daß $B = C$ nur für $\mu = 0$ und $\mu = 1$ gilt). Herr Birkhoff addiert nämlich in (4), (5), (6) zu unserem Ω die Konstante $\frac{1}{2}\mu(1-\mu)$, was bei Benutzung von bipolaren Koordinaten bequem ist.

Darwin[16]) hat, und zwar aus demselben Grunde, dieselbe additive Konstante, seine Jacobische Konstante, die wir mit D bezeichnen wollen (bei Darwin heißt sie C), weist aber von unserem C auch weitere Abweichungen auf. Denn zunächst hat Darwin andere Massen-, Zeit- und Längeneinheiten[17]) [übrigens legt er den Ursprung des (rotierenden) Koordinatensystems nicht in den Schwerpunkt, sondern in diejenige anziehende Masse, die links von der anderen liegt]. Die links liegende Masse ist bei

---

[14]) G. W. Hill, Coll. Math. Works **1** (1905), S. 284–335.
[15]) A. a. O. [8]), S. 269—270.
[16]) G. H. Darwin, Scientific Papers **4** (1911).
[17]) A. a. O. [16]), S. 3—4.

Darwin zehnmal größer als die andere, mit Rücksicht auf (3) haben wir also den Massenprozentsatz μ gleich $\frac{10}{11}$ [und nicht gleich $\frac{1}{11}$] zu setzen [obwohl freilich $\mu = \frac{1}{11}$ hier noch auf dasselbe hinauslaufen würde; vgl. (7)]. Um den Übergang von dem Darwinschen D zu unserem

$$C_{\frac{1}{11}} = C_{\frac{10}{11}} \tag{16}$$

zu finden, beachte man zunächst, daß jedenfalls

$$C_{\frac{10}{11}} = \mathfrak{A} D + \mathfrak{B} \tag{17}$$

gilt, wobei $\mathfrak{A}$ und $\mathfrak{B}$ numerische Konstanten bedeuten, die durch die Einheiten und durch den Nullpunkt des Jacobischen Energieniveaus bestimmt sind. $\mathfrak{A}$ und $\mathfrak{B}$ kann man nun folgendermaßen bestimmen. Das Jacobische Integral hat bei Darwin[16]) die Gestalt

$$V^2 - 10 \cdot \left(r^2 + \frac{2}{r}\right) - 1 \cdot \left(\varrho^2 + \frac{2}{\varrho}\right) + D = 0,$$

wobei V die Geschwindigkeit, r und ϱ die bipolaren Koordinaten bedeuten. Vergleicht man das mit (5) und (6) und läßt man D und daher auch das zugehörige $C = C_{\frac{10}{11}}$ über alle Grenzen wachsen, so folgt durch Dividieren der beiden Jacobischen Gleichungen ohne Mühe, daß bei diesem Grenzübergang

$$\frac{D}{C_{\frac{10}{11}}} \to \frac{10+1}{\mu+(1-\mu)} = 11 \qquad \left(\mu = \frac{10}{11}\right)$$

gilt, d. h. daß der Proportionalitätsfaktor in (17)

$$\mathfrak{A} = \frac{1}{11} \tag{17\alpha}$$

sein muß. Andererseits ist die Jacobische Konstante der äquilateralen Gleichgewichtslösungen bei Darwin durch $3(10+1) = 33$, bei uns aber durch (12a) [s. S. 355] gegeben, mit Rücksicht auf (17) muß also

$$3 - \frac{10}{121} = \mathfrak{A} \cdot 33 + \mathfrak{B},$$

d. h. wegen (17α)

$$\mathfrak{B} = -\frac{10}{121} \tag{17\beta}$$

sein. Durch Einsetzen von (17α) und (17β) in (17) folgt die gesuchte Reduktionsformel

$$C_{\frac{10}{11}} = \frac{1}{11} D - \frac{10}{121}. \tag{18}$$

In den Kopenhagener Rechnungen, die sich auf den Fall gleicher Massen ($\mu = \frac{1}{2}$) beziehen, sind die folgenden Gleichungen angegeben[18]):

[18]) E. Strömgren, Verhandlungen des Fünften skandinavischen mathematischen Kongresses (1922) in Helsingfors, Publ. Købn. Observ. Nr. 39, insb. S. 12.

$$\ddot{\xi} - 2\dot{\eta} = \frac{\partial \Omega}{\partial \xi}, \qquad \ddot{\eta} + 2\dot{\xi} = \frac{\partial \Omega}{\partial \eta}, \tag{19}$$

$$2\Omega = (\xi^2 + \eta^2) + \frac{8}{|\sqrt{(\xi+1)^2+\eta^2}|} + \frac{8}{|\sqrt{(\xi-1)^2+\eta^2}|}, \tag{20}$$

$$\dot{\xi}^2 + \dot{\eta}^2 = 2\Omega - K. \tag{21}$$

Der Vergleich mit (4), (5), (6), wobei $\mu = \frac{1}{2}$ zu setzen ist, zeigt unmittelbar, daß es sich um dasselbe Koordinatensystem wie bei uns handelt und daß zu unserem $C_{\frac{1}{2}}$ beim Übergang zu K diesmal keine additive Konstante, sondern nur ein durch die Verschiedenheit der Einheiten bedingter Proportionalitätsfaktor hinzukommt:

$$C_{\frac{1}{2}} = AK. \tag{22}$$

Die beiden gleichen Massen liegen nach (22) in den Punkten

$$\xi = \pm 1, \qquad \eta = 0. \tag{23}$$

Die Koordinaten der äquilateralen Librationspunkte sind daher $\xi = 0$, $\eta = \pm\sqrt{3}$, und (20) nimmt in diesen Punkten den Wert

$$3 + \frac{8}{|\sqrt{1+3}|} + \frac{8}{|\sqrt{1+3}|}$$

an. D. h., der K-Wert der beiden äquilateralen Gleichgewichtslösungen ist 11, ihr $C_{\frac{1}{2}}$-Wert aber nach (12b) viermal so klein, so daß in (22) der Proportionalitätsfaktor A gleich $\frac{1}{4}$ sein muß:

$$C_{\frac{1}{2}} = \frac{K}{4}. \tag{24}$$

Zur Kontrolle beachte man, daß der Wert von (20) in dem Librationspunkt $\xi = 0$, $\eta = 0$ gleich

$$0 + \frac{8}{|\sqrt{1}|} + \frac{8}{|\sqrt{1}|}$$

ist, also tatsächlich viermal größer als die Zahl (14)′. Man kann übrigens (24) auch wie folgt einsehen: Die Kopenhagener Zeiteinheit ist mit Rücksicht auf (4), (19) mit der unsrigen identisch, die Kopenhagener Längeneinheit aber mit Rücksicht auf (3), (23) nur halb so groß wie die unsrige, also die Kopenhagener Einheit des Geschwindigkeitsquadrates viermal kleiner als die unsrige; nun hat aber die Jacobische Konstante die Dimension des Quadrates einer Geschwindigkeit.

Unterdrückt man in (18) und (24) die Zeiger von C, so folgt durch Elimination von C:

$$K = \frac{4}{11} D - \frac{40}{121}. \tag{25}$$

Der Sinn der Formel (25) ist offenbar der folgende: numerische K-Werte ($\mu = \frac{1}{2}$) und numerische D-Werte ($\mu = \frac{10}{11}$), die einem und demselben

numerischen Werte von C in (6) entsprechen, sind miteinander durch (25) verknüpft. Da Darwins Zahlen D von der Größenordnung 40 sind und meist nur zweistellig angegeben werden, so kann man für (25) ruhig

$$K = \frac{4}{11} D - \frac{1}{3} \tag{26}$$

schreiben, wenn man die einer Darwinschen Bahn korrespondierende Bahn des Kopenhagener Materials aussuchen will. Hingegen erhält man eine völlig falsche Identifizierung der Darwinschen Satelliten b, c mit den Bahnen der Strömgrenschen Gruppe g, wenn man[19])

$$K = \frac{4}{11} D - 1 \tag{27}$$

setzt. Denn einerseits liegen die überhaupt in Betracht kommenden D-Werte etwa zwischen 38 und 39, sie haben daher ungefähr die Variationsbreite 1, die also nur anderthalbmal so groß ist wie der durch Benutzung von (27) begangene Fehler $\frac{2}{3}$; andererseits ist die Entwicklung der Gruppe g eben in diesem Gebiete der Jacobischen Konstante besonders stürmisch[20]). Fischer-Petersen[21]),[22]) hatte die Formel (27) wohl für andere Zwecke angegeben; man erhält sie nur dann, wenn man den Umstand, daß K und D zu zwei verschiedenen μ-Werten gehören, außer acht läßt, was für manche Zwecke geradezu geboten, für die Anwendungen aber, die wir vorhaben, durchaus unzulässig ist.

§ 2.

Der a priori zulässige Phasenraum.

Zur Vorbereitung des allgemeinen Falles müssen wir in diesem Paragraphen den Grenzfall $\mu = 0$ des näheren diskutieren.

Für $\mu = 0$ geht (5) in

$$\Omega_0 \equiv \Omega_0(\xi, \eta) = \frac{1}{2}(\xi^2 + \eta^2) + \frac{1}{|\sqrt{\xi^2 + \eta^2}|}, \tag{28}$$

also das Differentialsystem (4) in

$$\ddot{\xi} - 2\dot{\eta} = \frac{\partial \Omega_0}{\partial \xi} \equiv \xi\left(1 - \frac{1}{|\sqrt{\xi^2 + \eta^2}|^3}\right), \quad \ddot{\eta} + 2\dot{\xi} = \frac{\partial \Omega_0}{\partial \eta} \equiv \eta\left(1 - \frac{1}{|\sqrt{\xi^2 + \eta^2}|^3}\right) \tag{29}$$

---

[19]) Die Formel (27) erhält man so: Zunächst folgt durch Vergleich der Jacobischen Integrale im Unendlichen (vgl. S. 357 oben) jedenfalls $K = \frac{4}{11} D + \mathfrak{C}$, wobei $\mathfrak{C}$ eine additive Konstante bedeutet. Will man nun, daß zu $D = 33$ der K-Wert $K = 11$ gehören soll (trotzdem die Massenverteilung nicht dieselbe ist!), so folgt $\mathfrak{C} = -1$.

[20]) Siehe a. a. O. [18]), S. 15—21.

[21]) J. Fischer-Petersen, Sirius (1917), Publ. Købn. Observ. Nr. 27, S. 19 (Fußnote).

[22]) Vgl. auch J. Fischer-Petersen, Astr. Nachr. 201 (1915), S. 385 ff., Publ. Købn. Observ. Nr. 22.

und das Jacobische Integral (6) in

$$\frac{1}{2}(\dot{\xi}^2+\dot{\eta}^2)-\frac{1}{2}(\xi^2+\eta^2)-\frac{1}{|\sqrt{\xi^2+\eta^2}|}=-\frac{1}{2}C, \tag{30}$$

endlich die Gleichung (8) der Nullgeschwindigkeitskurve in

$$2\,\Omega_0(X,Y)\equiv X^2+Y^2+\frac{1}{|\sqrt{X^2+Y^2}|}=C \tag{31}$$

über, wobei $\xi=X$, $\eta=Y$ einen Punkt der Nullgeschwindigkeitskurve bedeutet.

Mit Rücksicht auf (29), (9) ist ein Punkt (ξ,η) dann und nur dann Librationspunkt, wenn er auf dem Einheitskreise

$$\xi^2+\eta^2=1 \tag{32}$$

liegt. Während es also für $0<\mu<1$ nur fünf Librationspunkte gibt, ist jetzt jeder Punkt der „Jupiterbahn" (32) ein Librationspunkt, in welchem die Gravitationskraft der (konservativen) Zentrifugalkraft das Gleichgewicht hält (im Innern des Einheitskreises ist die Gravitation stärker, in dem Außengebiet schwächer als die Zentrifugalwirkung). Daß dem so sein muß, folgt schon daraus, daß die beiden äquilateralen Librationspunkte bereits für $\mu\neq 0$ auf der heliozentrischen Jupiterbahn (vom Radius 1) liegen und daß andererseits das resultierende Kraftfeld für $\mu=0$ in bezug auf die Sonne $\xi=0$, $\eta=0$ [vgl. (3), wobei $\mu=0$ zu setzen ist] radialsymmetrisch wird, indem mit Rücksicht auf (28)

$$\Omega_0(\xi,\eta)\equiv\Omega_0(|\sqrt{\xi^2+\eta^2}|,0) \tag{33}$$

gilt [da der Kreis (32) nur zwei Punkte auf der Syzygienachse $\eta=0$ hat, so fallen für $\mu=0$ zwei der drei kollinearen Librationspunkte zusammen, der dritte liegt in dem anderen Schnittpunkt von $\eta=0$ mit (32); diese Unsymmetrie ist selbstverständlich, die beiden Punkte $\xi=\pm 1$ der Syzygienachse sind doch trotz der radialen Symmetrie keinesfalls gleichwertig, da der eine, nämlich $\xi=-1$, die (allerdings verschwindende, aber z. B. die zu $\mu\to 0$ gehörige Grenzlage der äquilateralen Librationspunkte dennoch bestimmende) Jupitermasse enthält[23])].

Setzt man

$$\varrho=|\sqrt{X^2+Y^2}|, \tag{34}$$

$$\text{also } 2\,\Omega_0(X,Y)\equiv 2\,\Omega_0(\varrho,0)=\varrho^2+\frac{2}{\varrho},\quad \frac{\partial}{\partial\varrho}2\,\Omega_0(\varrho,0)=2\left(\varrho-\frac{1}{\varrho^2}\right),$$

so geht die Gleichung (31) in

$$\varrho^3-C\varrho+2=0 \tag{35}$$

[23]) Eine ausführliche Diskussion der Bewegung der Librationspunkte bei unbeschränkt variablem Massenprozentsatz wird demnächst Herr Monroe Martin in dem Amer. Journ. of Math. veröffentlichen.

über, d. h. die zu C gehörige Nullgeschwindigkeitskurve (sofern sie existiert) besteht aus denjenigen Kreislinien um die Sonne $\varrho = 0$, deren Radien durch die (eventuellen) positiven Wurzeln dieser Gleichung dritten Grades gegeben sind. Da in (35) das von ϱ unabhängige Glied $= 2 > 0$ ist, so muß die Anzahl der negativen Wurzeln

(35)′ entweder eins oder drei

sein. Die Diskriminante von (35) ist

$$-4(-C)^3 - 27 \cdot 2^2 = 4(C^3 - 27),$$

also $\gtrless 0$ für $C \gtrless 3$. Für $C < 3$ gibt es also nur eine reelle Wurzel, und diese muß nach (35)′ negativ sein; für $C = 3$ muß es eine mindestens zweifache Wurzel geben, und da $\varrho = 1$ eine Wurzel ist, so ist $\varrho = 1$ mit Rücksicht auf (35)′ eine genau zweifache Wurzel; endlich gibt es für $C > 3$ drei voneinander verschiedene reelle Wurzeln, darunter nach Fourier auch positive, also nach (35)′ genau zwei positive Wurzeln, die ϱ_1 und $\varrho_2 > \varrho_1$ heißen mögen und die, wie wir gleich [vgl. (36)] sehen werden, voneinander durch die Zahl $\varrho = 1$ getrennt sind. Für $C > 3$ besteht also die zu C gehörige Nullgeschwindigkeitskurve aus zwei, voneinander durch die Jupiterbahn (32) getrennten Kreislinien mit den Radien ϱ_1, ϱ_2, und diese Kreise rücken für $C \to 3 + 0$ in die Jupiterbahn (32), also auch ineinander, um dann für $C < 3$ ganz zu verschwinden (d. h. für $C < 3$ gibt es überhaupt keine Nullgeschwindigkeitskurve). — Man überblickt die Verhältnisse am besten an Hand der Plummerschen Fläche $\zeta = 2\Omega_0(\xi, \eta)$ [vgl. auch S. 355], die jetzt nach (34) eine Rotationsfläche $\zeta = 2\Omega_0(\varrho, 0)$ ist; aus (34) folgt ferner, daß sie bei $\varrho = 0$ und $\varrho = +\infty$ ins Unendliche ragt, sonst regulär ist und nur entlang des Kreises $\varrho = 1$ eine zu $\zeta = 0$ parallele Tangentenebene, also entlang des Kreises $\varrho = 1$ ein Minimum hat, das nach (34) gleich 3 ist; sie wird also durch die Ebene $\zeta = C$ überhaupt nicht getroffen oder aber berührt oder endlich entlang zweier Kreise $\varrho = \varrho_1$, $\varrho = \varrho_2$, wobei offenbar

(36) $$0 < \varrho_1 < 1 < \varrho_2$$

ist, geschnitten, je nachdem $C \lessgtr 3$ gewählt wird.

Bisher haben wir für die Integrationskonstanten des Differentialsystems (29) keinerlei Einschränkung eingeführt, so daß die Bahnen auch nicht beschränkte Bahnen sowie auch Stoßbahnen sein können. Um solche Bahnen nunmehr der Übersicht halber auszuschließen, gehen wir für einen Moment mittels (2) in das siderische Achsenkreuz (x, y) zurück, wodurch sich (29) freilich in

(29)′ $$\ddot{x} = -\frac{1}{x^2+y^2} \cdot \frac{x}{|\sqrt{x^2+y^2}|}, \qquad \ddot{y} = -\frac{1}{x^2+y^2} \cdot \frac{y}{|\sqrt{x^2+y^2}|},$$

das Jacobische Integral (30) aber in

$$(30)' \qquad \frac{1}{2}(\dot{x}^2+\dot{y}^2)-(x\dot{y}-y\dot{x})-\frac{1}{|\sqrt{x^2+y^2}|}=-\frac{1}{2}C$$

transformiert, was eine Kombination des Energieintegrals

$$(30\,\mathrm{a})' \qquad \frac{1}{2}(\dot{x}^2+\dot{y}^2)-\frac{1}{|\sqrt{x^2+y^2}|}=h$$

und des Flächenintegrals

$$(30\,\mathrm{b})' \qquad x\dot{y}-y\dot{x}=c$$

von (30)′ darstellt:

$$(37) \qquad C=2(c-h).$$

Da die halbe große Achse a des Kegelschnittes sich aus der Energiekonstante h bekanntlich zu

$$(38) \qquad a=-\frac{1}{2h}$$

bestimmt, so verläuft die Bewegung dann und nur dann ganz im Endlichen, wenn

$$(39_1) \qquad h<0$$

ist. Ferner ist die Bahn dann und nur dann eine Stoßbahn, wenn sie geradlinig ist, d. h. wenn die Flächenkonstante verschwindet, so daß die Nichtstoßbahnen durch

$$(39_2) \qquad c\neq 0$$

charakterisiert sind. Führt man mittels der Substitution

$$(40) \qquad h=-\frac{1}{2\alpha^2}, \qquad c=\alpha|\sqrt{1-e^2}|$$

an Stelle von h, c das Parameterpaar α, e ein, so ist das durch die Ungleichungen (39_1), (39_2) eingeschränkte (h,c)-Gebiet auf das Gebiet

$$(41) \qquad \mu=0: \quad -\infty<\alpha<+\infty, \quad 0\leqq e<1; \quad \alpha\neq 0$$

des neuen Parameterpaares offenbar eineindeutig bezogen. Wegen (38), (40) ist

$$(42) \qquad a=\alpha^2,$$

wobei α wegen (40) dasselbe Vorzeichen hat wie die Flächenkonstante, diese hat aber dasselbe Vorzeichen wie die durch das dritte Keplersche Gesetz

$$(43) \qquad n=\frac{1}{\alpha^3}$$

festgelegte siderische mittlere Bewegung n, so daß also

$$(44) \qquad \operatorname{sign}\alpha=\operatorname{sign}n=\operatorname{sign}c$$

positiv oder negativ ist, je nachdem die Bahn, die ein Kreis ($e = 0$) oder eine eigentliche Ellipse ($0 < e < 1$) ist, siderisch direkt oder siderisch retrograd durchlaufen wird [zu einer Stoßbahn ($c = 0$) gehört eindeutig die Exzentrizität $e = 1$, sofern die Stoßbahn eine zur Strecke degenerierte Ellipse ist: $h < 0$; zu $h = 0$, $c = 0$ bzw. $h > 0$, $c = 0$ gehören hingegen die geradlinigen Grenzfälle der parabolischen bzw. hyperbolischen Bewegung]. Die siderische Periode ist mit Rücksicht auf (43) gleich

$$\frac{2\pi}{n}; \tag{45}$$

sie ist unabhängig von der Exzentrizität e und wird, je nach dem siderischen Umlaufssinn, positiv oder negativ gewählt.

Um die Jacobische Konstante C mittels des neuen Parameterpaares α, e auszudrücken, muß man nur (40) in (37) eintragen. So folgt für $\frac{1}{\alpha}$ die kubische Gleichung

$$\left(\frac{1}{\alpha}\right)^3 - C \cdot \frac{1}{\alpha} + 2\left|\sqrt{1-e^2}\right| = 0 \tag{46}$$

mit der Diskriminante

$$-4(-C)^3 - 27\left(2\left|\sqrt{1-e^2}\right|\right)^2 = 4\left[C^3 - 27(1-e^2)\right], \tag{47}$$

die also

$$\gtreqless 0 \text{ ist, je nachdem } C \gtreqless 3\sqrt[3]{1-e^2} \tag{48}$$

gewählt wird. Da in (46) das von α unabhängige Glied wegen $0 \leqq e < 1$ positiv ist, so folgt genau so wie vorher bei (35), daß, wenn e irgendwie fest gewählt wird, bei jedem C genau eine negative Wurzel existiert und daß die beiden anderen Wurzeln positiv und voneinander verschieden oder aber nicht reell sind, je nachdem C größer oder kleiner ist als die jedenfalls positive Zahl $3\sqrt[3]{1-e^2}$, während zur verschwindenden Diskriminante ($C = 3\sqrt[3]{1-e^2}$) die positive Doppelwurzel $\sqrt[3]{1-e^2}$ gehört. Ferner können die beiden positiven Wurzeln (wobei also $C > 3\sqrt[3]{1-e^2}$ sein soll) nicht beide $< \sqrt[3]{1-e^2}$ oder beide $> 1 > \sqrt[3]{1-e^2}$ sein; die Behauptung ist nämlich nach (35), (36) im Falle $e = 0$ bei jedem $C(>3)$ richtig; wäre sie also bei einem Zahlenpaar $C = C'$, $e = e'$ unrichtig, so müßte es aus Stetigkeitsgründen ein e'' mit $0 < e'' \leqq e'$ geben derart, daß für $C = C'$, $e = e''$ beide positiven Wurzeln gleich $\sqrt[3]{1-(e'')^2}$ sind, es würde also eine Doppelwurzel geben, im Widerspruch damit, daß nach Voraussetzung $C > 3\sqrt[3]{1-e^2}$ gilt, also (47) nicht verschwindet. — Damit ist folgendes bewiesen:

e sei der Bedingung $0 \leqq e < 1$ gemäß beliebig fest gewählt. Ist $C > 3\sqrt[3]{1-e^2}$, so kann man die halbe große Achse a auf genau zwei

Weisen derart bestimmen, daß die Bewegung siderisch direkt ist und eine der beiden möglichen Halbachsen ist kleiner, die andere größer als $\sqrt[3]{1-e^2}$, sie fallen ferner für $C=3\sqrt[3]{1-e^2}$ miteinander, also auch mit $\sqrt[3]{1-e^2}$ zusammen, während für $C<3\sqrt[3]{1-e^2}$ jede siderische direkte Bewegung unmöglich ist; hingegen kann a bei jedem $C(-\infty<C<+\infty)$ auf genau eine Weise derart bestimmt werden, daß die Bewegung siderisch retrograd ist. — Damit sind alle den beiden Bedingungen (39), (40) genügenden Bewegungsformen mittels des Parameterpaares C, e eineindeutig charakterisiert (die Perihellänge der Epoche und die Epoche gehen uns hier nicht an, d. h. zwei Lösungen, die sich nur in Perihellänge oder Epoche unterscheiden, gelten als identisch). Die siderisch retrograden Lösungen (mit beliebigen Halbachsen) sind also auf die Punkte des Streifens

$$(49)\qquad \mathfrak{S}^-:\quad -\infty<C<+\infty,\quad 0\leqq e<1$$

der (C, e)-Ebene eineindeutig bezogen. Ferner entspricht jedem Punkte des Streifens

$$(50)\qquad \mathfrak{S}^+:\quad 3<C+\infty,\quad 0\leqq e<1$$

genau eine siderisch direkte Lösung mit einer Halbachse $a=a_1<\sqrt[3]{1-e^2}$ und genau eine von der ersten verschiedene siderisch direkte Lösung mit einer Halbachse $a=a_2>\sqrt[3]{1-e^2}$, wobei

$$(51\alpha)\qquad a_1=a_1(C,e),\quad a_2=a_2(C,e)\quad (a_1<\sqrt[3]{1-e^2}<a_2)$$

Funktionen des Ortes auf $\mathfrak{S}^+$ sind, so daß wir $\mathfrak{S}^+$ der Eindeutigkeit halber in zwei Exemplaren $\mathfrak{S}_1^+, \mathfrak{S}_2^+$ einführen wollen:

$$(51\beta)\quad \text{in } \mathfrak{S}_1^+:\ a=a_1;\quad \text{in } \mathfrak{S}_2^+:\ a=a_2;\quad a_1(C,e)\neq a_2(C,e),$$

so daß wir insgesamt drei „Phasenstreifen“ $\mathfrak{S}^-, \mathfrak{S}_1^+, \mathfrak{S}_2^+$ haben [$\mathfrak{S}_1^+$ oder $\mathfrak{S}_2^+$ gelten also trotz (50), (49) keinesfalls als Teilmengen von $\mathfrak{S}^-$]. $\mathfrak{S}_1^+ + \mathfrak{S}_2^+$ ergibt noch nicht alle siderisch direkten Bahnen, da doch unter (50) nicht nur $C\geqq 3\sqrt[3]{1-e^2}$, sondern $C>3$ vorausgesetzt ist, doch ist die Annahme $C>3$ notwendig, wenn man will, daß bei jedem C auch für kleine e $(\geqq 0)$ zwei siderisch direkte Bahnen existieren; übrigens haben die siderisch direkten Bahnen, für welche nicht bereits $C>3$ gilt, d. h. die nicht in $\mathfrak{S}_1^+ + \mathfrak{S}_2^+$ enthalten sind, für die Anwendungen sowieso keine Bedeutung. — Ist $C'>3$, so ist $C=C'$ eine auf der die Kreisbahnen enthaltenden Gerade $e=0$ senkrecht stehende Strecke in einem jeden der drei Phasenstreifen, in $\mathfrak{S}^-$ freilich auch ohne die Annahme $C'>3$. Wir nennen diese Strecke kurz das *C'-Lot* des jeweiligen Phasenstreifens, und den Punkt $e=0$ des Lotes, der die zu C' gehörige Kreisbahn des Phasenstreifens darstellt, den *Anfangspunkt* des Lotes, endlich die Gesamtheit der

Lotanfangspunkte, d. h. die Gerade $e = 0$ der Kreisbahnen, kurz die *Randgerade* des Phasenstreifens. Das C-Lot enthält also alle Bahnen des Phasenstreifens, die eine feste Jacobische Konstante haben. α, a und n sind bei festem C stetige Funktion von e, d. h. des Ortes auf dem C-Lot. Dasselbe gilt mit Rücksicht auf (45) auch von der siderischen Periode. Ganz anders steht aber die Sache bei der synodischen Periode, zu der wir nun übergehen. Was die Variabilität der siderischen Periode (45) bzw. von (43) auf irgendeinem Lote anlangt, so gilt übrigens nach (46)

$$\frac{3}{\alpha^2}\frac{\partial\frac{1}{\alpha}}{\partial e} - C\frac{\partial\frac{1}{\alpha}}{\partial e} - \frac{2e}{|\sqrt{1-e^2}|} = 0,$$

oder, da nach (46)

$$C = \frac{1}{\alpha^2} + 2\alpha|\sqrt{1-e^2}| \tag{52}$$

ist,

$$\frac{\partial\frac{1}{\alpha}}{\partial e} = \frac{e}{\sqrt{1-e^2}}\cdot\frac{\alpha^2}{1-\alpha^3|\sqrt{1-e^2}|} \qquad (C \text{ fest}). \tag{53}$$

Bei der Frage nach der Periodizität in dem synodischen Achsenkreuz (ξ, η) sind nun zwei Fälle zu unterscheiden, je nachdem I. $e = 0$ oder II. $e \neq 0$ ist, und im Fall II wieder zwei Unterfälle, je nachdem n II_1. irrational oder II_2. rational, also gleich $p:q$ ist, wobei die ganzen Zahlen $p > 0$, q teilfremd und damit eindeutig bestimmt sind. Im Falle I ist die Bewegung auch in dem rotierenden Achsenkreuz (ξ, η) ausnahmslos (d. h. auch bei irrationalem n) periodisch mit der primitiven Periode

$$e = 0: \quad T = \frac{2\pi}{n-1} = 2\pi m, \tag{54}$$

wobei

$$m = \frac{1}{n-1} \tag{55}$$

der „Hillsche Periodenquotient" ist; im Falle II_1 ist die Bewegung in dem synodischen Achsenkreuz überhaupt nicht periodisch, im Falle II_2 endlich periodisch mit der primitiven Periode

$$e \neq 0: \quad \tau = \pm 2\pi q, \tag{56}$$

wofür man wegen

$$n = \frac{p}{q}; \qquad p > 0,\ q \gtrless 0,\ (p, q) = 1 \tag{57}$$

auch

$$\tau = \pm\frac{2\pi}{n}\cdot p \tag{56'}$$

schreiben kann; die synodischen primitiven Perioden (54), (56) sind dabei positiv oder negativ angesetzt, je nachdem die Bewegung synodisch direkt

oder synodisch retrograd ist [d. h. die Vorzeichen sind in (56), (56)′ passend zu wählen; in (54) stimmt das Vorzeichen von vornherein], was übrigens für $e \neq 0$ auf eine von t unabhängige Weise nicht von vornherein definiert ist. Die allgemein gültige siderische Periode (45) stimmt dem Betrage nach mit der zu $e = 0$ gehörigen synodischen Periode (54) nie, mit der zu $e \neq 0$ gehörigen synodischen Periode (56) nur für ganzzahlige Werte von n^{-1} überein [wegen (56)′, (57)]. — Im Falle (57) geht die (auch für irrationale n gültige) Formel (54) in

$$(54)' \qquad e = 0: \quad T = \frac{2\pi q}{p - q} \qquad \left(n = \frac{p}{q}\right)$$

über, so daß man die (nur für rationale n vorhandene) Formel (56) auch in der Gestalt

$$(56)^* \qquad \tau = (p - q)\, T$$

schreiben kann. D. h., die zu $e = 0$ gehörige synodische primitive Periode ist stets ein Teiler der zu $e \neq 0$ gehörigen synodischen primitiven Periode, und zwar dann und nur dann kein *echter* Teiler, wenn $p - q = \pm 1$ ist. Da nach (55), (57)

$$(55)' \qquad m = \frac{q}{p - q} \qquad \left(n = \frac{p}{q}\right)$$

gilt, so ist daher τ nur für ganzzahlige Werte von m gleich T, für gebrochene Werte von m aber ein Multiplum von T, und zwar ein desto größeres Multiplum, je näher der Bruch n, oder — was nach (55) dasselbe ist — der Bruch m an eine Irrationalzahl heranrückt, d. h. bei Annäherung an eine Irrationalzahl wird τ bei beliebig kleinem e unendlich groß, während T beschränkt bleibt; vgl. (54), (56). — Hat m einen von 0 und -1 verschiedenen ganzzahligen Wert, so wollen wir aus später (S. 388) ersichtlich werdenden Gründen sagen, m sei ein Lückenzentrum.

In den Phasenstreifen (vgl. S. 364) verhält sich also die Sache wie folgt. Nennt man einen Punkt (C, e) des Streifens *a priori zulässig*, wenn der zu (C, e) gehörige, d. h. gemäß (43), (46) bestimmte Wert von n rational, d. h. wenn die zu (C, e) gehörige Lösung auch in dem synodischen Achsenkreuz periodisch ist, so besteht die a priori zulässige Mannigfaltigkeit nicht aus dem ganzen Streifen, sondern nur aus abzählbar vielen, in dem Streifen überall dicht liegenden analytischen Kurven, nämlich aus den zu einem rationalen n_0 gehörigen synchronen Kurven und aus der Randgeraden $e = 0$ [n ist nach (43), (46) eine Funktion $n(C, e)$ des Ortes im Streifen, und unter der zu n_0 gehörigen synchronen Kurve ist selbstverständlich die Kurve $n(C, e) = n_0$ verstanden; vgl. übrigens S. 372 weiter unten]. Die primitive synodische Periode ist eine Funktion des Ortes in der a priori zulässigen Mannigfaltigkeit, und zwar gleich $\tau = \tau(C, e) = \pm 2\pi q$ oder

gleich $T = T(C, 0) = 2\pi m$, je nachdem der Punkt (C, e) im Innern eines Lotes $(e \neq 0)$ oder auf der Randgerade $(e = 0)$ des Streifens liegt. Auf jedem Lote (C fest) ist $\tau = \tau(C, e)$ nur auf einer überall dichten, abzählbaren Punktmenge (nämlich in den Schnittpunkten des Lotes mit den zu irgendeinem rationalen n gehörigen synchronen Kurven) erklärt. Ferner ist $\tau = \tau(C, e)$ auf einem Lote in der Umgebung keines Punktes des Lotes beschränkt; insbesondere gilt also in jedem der drei Streifen

$$\limsup_{e \to +0} |\tau(C, e)| \equiv +\infty, \tag{58}$$

wobei der Grenzübergang darin besteht, daß der a priori zulässige Punkt des Phasenstreifens entlang des zu C gehörigen Lotes gegen die Randgerade konvergiert. — Es sei hier wiederholt (vgl. S. 364), daß die Phasenstreifen nur zweidimensional sind, so daß zwei der vier Integrationskonstanten, und zwar die Epoche und die dazu gehörige Perihellänge, die übrigens für $e = 0$ unbestimmt wird, noch fehlen; doch ist die durch diese beiden noch verfügbaren Integrationskonstanten gegebene Freiheit für das Nachfolgende auch in dem synodischen Achsenkreuz belanglos bzw. ohne Wert, so daß wir festsetzen wollen, daß die Epoche einfach $t = 0$ ist und daß das Passieren des Perihelpunktes auf der positiven ξ-Halbachse, und zwar zur Zeit $t = 0$ stattfindet[24]).

§ 3.

Nähere Bestimmung der a priori zulässigen Gebiete.

In diesem Paragraphen wollen wir die a priori zulässigen Gebiete der Phasenstreifen, also zunächst (bis S. 372) den Rand $e = 0$ und dann (S. 372—373) die Folge der synchronen Kurven des näheren analysieren.

* * *

Die zu $e = 0$ gehörige a priori zulässige Mannigfaltigkeit, d. h. das Kontinuum der orientierten Keplerschen Kreisbahnen, ist nach S. 362 und mit Rücksicht auf (41) eineindeutig bezogen auf die lineare Mannigfaltigkeit

$$-\infty < \alpha < +\infty, \tag{59}$$

aus welcher die Stelle $\alpha = 0$ auszulassen ist, da nach (42) zugleich mit α auch der Bahnradius $a = \alpha^2$ [vgl. (42)] verschwindet, so daß die Bahn auf die Sonne zusammenschrumpft. Ferner wollen wir auch $\alpha = 1$ und $\alpha = -1$ ausschließen, da a beidemal $= 1$ wird, so daß die Keplersche Kreisbahn des dritten Körpers auf der siderisch direkt bzw. siderisch retrograd durch-

[24]) Über die Schwarzschildschen Untersuchungen [vgl. z. B. a. a. O. [7])] vgl. S. 401 weiter unten.

laufenen Jupiterbahn (32) liegt; zwar ist jetzt die Jupitermasse μ gleich Null, wir wissen aber (S. 360), daß im Falle $\alpha^2 = a = 1$ für $\mu = 0$ und für $0 < \mu \to 0$ voneinander durchaus verschiedene Vorkommnisse zu erwarten sind; dementsprechend soll die Liberationsgenealogie nicht in den folgenden Paragraphen dieser Arbeit, sondern in einer späteren Mitteilung aufgebaut werden. — Durch die ausgeschlossenen Stellen

$$\alpha = 0, \quad \alpha = 1, \quad \alpha = -1, \quad \alpha = \pm\infty, \tag{60}$$

zu welchen nach (43) die n-Werte

$$n = \pm\infty, \quad n = 1, \quad n = -1, \quad n = 0, \tag{61}$$

also nach (55) die m-Werte

$$m = 0, \quad m = \pm\infty, \quad m = -\frac{1}{2}, \quad m = -1, \tag{62}$$

endlich nach (52) $[e = 0]$ die C-Werte

$$C = +\infty, \quad C = -1, \quad C = 3 \tag{63}$$

gehören, wird das Kontinuum (59) in vier getrennte Stücke geteilt, die wir in der Reihenfolge

$$\begin{aligned} &\mathbf{1}: && 0 < \alpha < 1, \\ &\mathbf{2}: && 1 < \alpha < +\infty, \\ &\mathbf{3}: && -\infty < \alpha < -1, \\ &\mathbf{4}: && -1 < \alpha < 0 \end{aligned} \tag{64}$$

mit **1**, **2**, **3**, **4** bezeichnen wollen (vgl. Zeile I der Tabelle). Da die Relationen (43), (55) nach Ausschluß der Stellen (60), (61), (62) im reellen Gebiete eindeutig umkehrbar sind:

$$n = \frac{1}{\alpha^3}, \qquad \alpha = \frac{1}{\sqrt[3]{n}}, \tag{65}$$

$$m = \frac{1}{n-1}, \quad n = 1 + \frac{1}{m}, \tag{66}$$

so ist (59) auf

$$-\infty < n < +\infty \tag{67}$$

und auch auf

$$-\infty < m < +\infty \tag{68}$$

nach Ausschluß von (60), (61), (62) eineindeutig bezogen, so daß (67) bzw. (68) durch (61) bzw. (62) in vier getrennte Gebiete geteilt wird, auf welche die vier Gebiete (64) durch (65) bzw. (66) eineindeutig und stetig, also auf eine (innerhalb eines der vier Gebiete) monotone Weise abgebildet werden (vgl. II und III in der Tabelle, wobei korrespondierende Parameterwerte untereinander geschrieben sind, so daß auch der Durchlaufssinn der vier Gebiete ersichtlich ist).

		1	2	3	4	Zweiteilung von 4		
						4_1	Grenze	4_2
I	$\alpha = \sqrt{a}$	$0 < \alpha < 1$	$1 < \alpha < +\infty$	$-\infty < \alpha < -1$	$-1 < \alpha < 0$	$-1 < \alpha < -\frac{1}{2}$	$\alpha = -\frac{1}{2}$	$-\frac{1}{2} < \alpha < 0$
II	n	$+\infty > n > 1$	$1 > n > 0$	$0 > n > -1$	$-1 > n > -\infty$	$-1 > n < -8$	$n = -8$	$-8 > n > -\infty$
III	m	$0 < m < +\infty$	$-\infty < m < -1$	$-1 < m < -\frac{1}{2}$	$-\frac{1}{2} < m < 0$	$-\frac{1}{2} < m < -\frac{1}{9}$	$m = -\frac{1}{9}$	$-\frac{1}{9} < m < 0$
IV	$j =$ krit. m	$1, 2, 3, \ldots$	$\ldots, -4, -3, -2$	*	*			
V	krit. n	$\frac{2}{1}, \frac{3}{2}, \frac{4}{3}, \cdots$	$\cdots, \frac{3}{4}, \frac{2}{3}, \frac{1}{2}$	*	*			
VI	krit. a	$\left\lvert\sqrt[3]{\frac{1}{2}}\right\rvert^2, \left\lvert\sqrt[3]{\frac{2}{3}}\right\rvert^2, \left\lvert\sqrt[3]{\frac{3}{4}}\right\rvert^2, \ldots$	$\ldots, \left\lvert\sqrt[3]{\frac{4}{3}}\right\rvert^2, \left\lvert\sqrt[3]{\frac{3}{2}}\right\rvert^2, \left\lvert\sqrt[3]{\frac{2}{1}}\right\rvert^2$	*	*			
VII	Häufungsstelle der krit. a	$a = 1 - 0$	$a = 1 + 0$	*	*			
VIII	C	$+\infty > C > 3$	$3 < C < +\infty$	$-\infty < C < -1$	$-1 < C < +\infty$	$-1 < C < 3$	$C = 3$	$3 < C < +\infty$
IX	Erstes krit. C	$(C)_{m=1} = 3{,}1748$	$(C)_{m=-2} = 3{,}1498$	*	*			
X	Nullgeschwindigkeitskreis	äußerer: $a < \varrho < 1$	innerer: $a > \varrho > 1$	* *	s. Zweiteilung	* *	$a = \varrho$	äußerer: $a < \varrho < 1$
XI	ϱ	$0 < \varrho \equiv \frac{1}{\alpha_2} < 1$	$1 < \varrho \equiv \frac{1}{\alpha_1} < +\infty$	* *	s. Zweiteilung	* *	$\varrho = 1$	$1 > \varrho \equiv \frac{1}{\alpha_2} > 0$
XII	sign n	+	+	−	−			
XIII	sign m	+	−	−	−			
XIV	Gruppenanfang	$m = +0$	$m = -1 - 0$	$m = -1 + 0$	$m = -0$			
XV	Gruppe	g	l	m	f			

Nicht so einfach liegen die Dinge bei dem Parameter C, da die aus (52) für den vorliegenden Fall $e = 0$ folgende Beziehung

$$e = 0: \qquad C = 2\alpha + \frac{1}{\alpha^2}, \tag{69}$$

die mit Rücksicht auf (65), (66) auch in der Gestalt

$$C = \frac{3 + \frac{1}{m}}{\sqrt[3]{1 + \frac{1}{m}}} \tag{70}$$

geschrieben werden kann, das Gebiet (59) auf das ganze Gebiet

$$-\infty < C < +\infty \tag{71}$$

nach Ausschluß von (60), (63) zwar offenbar eindeutig, aber nicht eineindeutig bezieht. So viel ist allerdings richtig, daß C innerhalb eines jeden der vier Gebiete (64) [aber, im Gegensatz zu n oder m, nicht auch innerhalb der Vereinigung (59) dieser Gebiete] eine eindeutig umkehrbare Funktion von α ist, da doch nach (69) die Ableitung

$$\frac{dC}{d\alpha} = 2\left(1 - \frac{1}{\alpha^3}\right) \tag{72}$$

in keinem der vier offenen Gebiete (64) verschwinden oder unendlich werden kann, so daß C innerhalb eines jeden der vier Gebiete (64) eine (nirgends konstante) monotone Funktion von α sein muß. Was aber das Vollgebiet (64) [nach Ausschluß von (60)] anlangt, so verhält es sich mit Rücksicht auf die korrespondierenden Parameterwerte (60), (63) so, daß **1** dem Gebiete

$$3 < C < +\infty, \tag{73}$$

aber **2** demselben Gebiete (73) und **4** dem die Halbgerade (73) ebenfalls enthaltenden Gebiete

$$-1 < C < +\infty \tag{74}$$

eineindeutig entspricht, während **3** auf das zu (74) komplementäre Gebiet

$$-\infty < C < -1 \tag{75}$$

eineindeutig bezogen ist (vgl. VIII in der Tabelle, wobei wieder auch der Durchlaufssinn angegeben ist). Läuft also C von 3 bis $+\infty$, so erhält man **1** vollständig, aber auch **2** vollständig, und außerdem wegen (74) einen Teil von **4**, während zu

$$-1 < C < 3 \tag{76}$$

nur der andere Teil von **4** (vgl. die Zweiteilung von **4** in der Tabelle), endlich zu (75) nur **3** gehört, so daß es zu C drei orientierte Kreisbahnen gibt oder nur eine einzige, je nachdem $C \gtrless 3$ ist. Dem entspricht es, daß

(69) nichts weiter ist als die kubische Gleichung

$$\left(\frac{1}{\alpha}\right)^3 - C\cdot\frac{1}{\alpha} + 2 = 0, \tag{77}$$

deren Diskriminante $4(C^3 - 27)$ für $C > 3$ positiv, für $C < 3$ negativ wird [vgl. (47), (48)]. Mit Beachtung des auf S. 362 bis 364 Gesagten sowie der Zeilen I und VIII der Tabelle erkennen wir jetzt, daß **1** den Rand des Phasenstreifens $\mathfrak{S}_1^+$, **2** den Rand von $\mathfrak{S}_2^+$, endlich die Vereinigung von **3** und **4** den Rand von $\mathfrak{S}^-$ bildet:

$$\mathbf{1}\colon\ \mathfrak{S}_1^+, \qquad \mathbf{2}\colon\ \mathfrak{S}_2^+, \qquad \mathbf{3}+\mathbf{4}\colon\ \mathfrak{S}^-. \tag{78}$$

Nach S. 363, wobei $e = 0$ zu setzen ist, oder nach den Zeilen I, VIII der Tabelle hat (77) bei jedem reellen C genau eine negative Wurzel

$$\frac{1}{\alpha'} < 0 \qquad (-\infty < C < +\infty), \tag{79}$$

für $C < 3$ keine weitere reelle Wurzel, für $C > 3$ aber zwei positive Wurzeln $\frac{1}{\alpha_1}$, $\frac{1}{\alpha_2}$, die voneinander durch die Zahl 1 getrennt sind und die wir (damit α_1 in **1** und α_2 in **2** liegt) derart beziffern wollen, daß

$$\frac{1}{\alpha_1} > 1 > \frac{1}{\alpha_2} > 0 \tag{80}$$

wird. Vergleicht man (35), (36) mit (77), (80), so folgt

$$\varrho_1 = \frac{1}{\alpha_2}, \qquad \varrho_2 = \frac{1}{\alpha_1} \tag{81}$$

[vgl. X, XI in der Tabelle; in **3** und $\mathbf{4}_1$ gibt es keinen Nullgeschwindigkeitskreis, da $C < 3$ ist, während doch ϱ, als Radius, > 0 sein muß]. Ist $C > 3$ gewählt, so muß die negative Wurzel (79) mit Rücksicht auf die Zeilen I, VIII der Tabelle zu $\mathbf{4}_2$ gehören. Gehören also $\alpha_1, \alpha_2, \alpha'$ zu einem und demselben Werte von $C > 3$, so muß mit Rücksicht auf die Zeile I der Tabelle

$$0 < \alpha_1 < 1, \qquad 1 < \alpha_2 < +\infty, \qquad -\frac{1}{2} < \alpha' < 0 \tag{82}$$

sein, so daß die siderischen mittleren Bewegungen

$$n_1 = n_1(C) = \frac{1}{\alpha_1^3}, \qquad n_2 = n_2(C) = \frac{1}{\alpha_2^3}, \qquad n' = n'(C) = \frac{1}{\alpha'^3} \tag{83}$$

den Ungleichungen

$$n' < -8 < 0 < n_2 < 1 < n_1 \qquad (C > 3) \tag{84}$$

genügen.

Die Bahn heißt eine untere oder eine obere Bahn, je nachdem ihr Radius $a = \alpha^2$ kleiner oder größer als eins ist, d. h. je nachdem sie im Inneren oder im Äußeren der Jupiterbahn verläuft, also je nachdem

die Bewegung synodisch direkt oder synodisch retrograd ist. XII, XIII folgen aus II, III und besagen (vgl. § 2), daß der siderische Umlaufssinn bei den siderisch direkten unteren Bahnen (1) sowie bei allen [d. h. sowohl bei den oberen (3) als auch bei den unteren (4)] siderisch retrograden Bahnen mit dem synodischen Umlaufssinn übereinstimmt, daß aber die siderisch direkten oberen Bahnen (2) synodisch retrograd sind. Es sei hier erwähnt, daß n nach (66) für $|m| \geqq 1$ gewiß $\geqq 0$, also, da $n = 0$ nach (65) sinnlos ist, sogar > 0 ist. Mit Rücksicht auf (65) gilt daher

$$\alpha > 0 \quad \text{für} \quad |m| \geqq 1. \tag{85}$$

Nach (65), (66) gilt allgemein

$$n = \frac{m+1}{m}, \qquad \alpha = \sqrt[3]{\frac{m}{m+1}} \qquad (\alpha = \pm\sqrt{a},\ a > 0). \tag{86}$$

In der Zeile IV der Tabelle sind die sogenannten kritischen m-Werte:

$$m = j; \quad j = 1, \pm 2, \pm 3, \ldots \tag{87}$$

(vgl. S. 366 oben und S. 380 weiter unten) angegeben [die beiden Werte

$$m = 0, \quad m = -1 \tag{88}$$

kommen nicht vor, da $m = 0$ und $m = -1$ bei (62) ausgeschlossen worden sind]. Mit Rücksicht auf (85) können nur die beiden siderisch direkten Spalten 1, 2 kritische m-Werte (87) enthalten. V und VI folgen wegen (86) aus IV, und VII folgt aus VI. Was endlich IX anlangt, so genügt zu bemerken, daß nach (70)

$$3{,}1748 = \sqrt[3]{32} = (C)_{m=1} > (C)_{m=2} > (C)_{m=3} \cdots \to 3; \quad (C)_{m=0} = +\infty, \tag{89}$$

$$3{,}1498 = \frac{5}{2}\sqrt[3]{2} = (C)_{m=-2} > (C)_{m=-3} > (C)_{m=-4} \cdots \to 3 \tag{90}$$

gilt. Auf XIV und XV kommen wir weiter unten zurück.

* * *

Bisher handelte es sich um den zu $e = 0$ gehörigen Teil der a priori zulässigen Mannigfaltigkeit. In einem Phasenstreifen (C, e) ist α oder n eine Funktion des Ortes, und die Gesamtheit derjenigen Punkte (C, e), für welche $n = n(C, e)$ einen festen Wert hat, haben wir eine synchrone Kurve genannt. Der zu $e \neq 0$ gehörige Teil der a priori zulässigen Mannigfaltigkeit besteht aus der Folge derjenigen synchronen Kurven, für welche $n = n(C, e) =$ konst. eine Rationalzahl n ist. Die Gleichung der zur festen Rationalzahl n gehörigen synchronen Kurve des (C, e)-Streifens ist nach (43), (46)

$$n - C\sqrt[3]{n} + 2|\sqrt{1 - e^2}| = 0 \qquad (n = \text{feste Rationalzahl}). \tag{91}$$

Die synchronen Kurven sind also Ellipsenbogen, die von denjenigen (überall dicht liegenden) Punkten C der Randgeraden $e = 0$ des Phasenstreifens (C, e) ausgehen, für welche

$$(91') \qquad e = 0: \qquad C = \frac{n+2}{\sqrt[3]{n}} \qquad (n = \text{Rationalzahl})$$

gilt. Von dem Punkte (91') geht nur ein Ellipsenbogen aus, d. h. die Kurve (91) wird durch die Randgerade $e = 0$ nicht berührt. Die Folge (91) der synchronen Kurven durchquert den Phasenstreifen überall dicht, d. h. jedes e-Lot $C = \text{konst.}$ wird durch die Folge der synchronen Kurven überall dicht geschnitten, da doch $n = n(C, e)$ bei festem C eine nicht konstante analytische Funktion von e ist, die also auf einer überall dichten e-Menge rational werden muß (wobei freilich (58) nicht zu vergessen ist).

§ 4.

Die Sortenmannigfaltigkeit.

Wir können nunmehr das Poincarésche *Sorten*problem[3]), [5]), das nur ein periodisches *Störungs*problem[12]), [1]) ist, exakt und allgemein formulieren[25]); weshalb ich sage, es handle sich *nur* um ein Störungsproblem, ist dem Numeriker von vornherein klar[12]), wird aber auf S. 389 ff. auch mathematisch begründet, und zwar wird es sich dabei nicht um eine nur technisch-numerische, sondern um eine rein funktionentheoretische Frage handeln.

Es sei (C, e) ein fest gewählter a priori zulässiger Punkt eines der drei Phasenstreifen $\mathfrak{S}_1^+$, $\mathfrak{S}_2^+$, $\mathfrak{S}^-$.

Es sei zunächst vorausgesetzt, daß es eine positive Zahl $M = M(C, e)$ gibt, für welche die folgenden Voraussetzungen erfüllt sind: Erstens gibt es für

$$(92) \qquad 0 < \mu < M(C, e)$$

eine von dem Scharparameter μ abhängige stetige Schar von periodischen Lösungen

$$(93) \qquad \xi = \sum_{k=-\infty}^{+\infty} A_k(\mu) \exp \frac{2\pi k i t}{T(\mu)}, \qquad \eta = \sum_{k=-\infty}^{+\infty} B_k(\mu) \exp \frac{2\pi k i t}{T(\mu)}$$

---

[25]) Dabei wollen wir die auf S. 354 eingeführte Festsetzung über die Bezeichnung der beiden Massen μ, $1-\mu$, die aber miteinander vertauscht werden könnten, beibehalten. In Wirklichkeit handelt es sich freilich stets um *zwei* Sortenmannigfaltigkeiten, die eine gehört zu $(\mu, 1-\mu)$, die andere zu $(1-\mu, \mu)$. Ihre gegenseitige Beziehung wird mit Rücksicht auf die auf S. 357 gemachte Bemerkung über die Darwinschen Planeten ohne weiteres einleuchten. — Wie gesagt, es handelt sich im Texte nach wie vor nur um die auf S. 354 festgelegte Anordnung und nicht auch um die andere, die ebenso behandelt werden könnte.

von (4), die für $\mu \to 0$ in die gegebene, durch den Punkt (C, e) des Streifens charakterisierte Lösung von (29) stetig übergeht; zweitens ist (93) reell:

$$A_k = \bar{A}_{-k}, \quad B_k = \bar{B}_{-k}, \quad T \gtrless 0; \tag{94}$$

drittens wird durch (93) für $t = 0$ ein auf der positiven ξ-Halbachse gelegener Punkt dargestellt:

$$(\xi)_{t=0} > 0, \quad (\eta)_{t=0} = 0; \tag{95}$$

viertens sind die in dem Intervall (92) stetigen Funktionen

$$A_k = A_k(\mu), \quad B_k = B_k(\mu), \quad T = T(\mu) \tag{96}$$

analytisch für

$$0 \leqq \mu < M(C, e) \tag{97}$$

[sie brauchen dabei nicht regulär zu sein; ist $\mu_0 \geqq 0$ ein Verzweigungspunkt in (97), so wähle man für $\mu > \mu_0$ irgendeine der möglichen Bestimmungen]; fünftens ist die Jacobische Konstante der Lösung (93) unabhängig von μ, also gleich der gegebenen, zu $\mu = 0$ gehörigen Zahl C (dies ist nur eine Art Normierung; vgl. weiter unten); sechstens sei $T(\mu)$ für $\mu > 0$ (also nicht notwendig auch für $\mu = 0$) die primitive Periode der Schar für „hinreichend kleines" μ. — Gibt es zu (C, e) eine Zahl $M(C, e)$ von dieser Beschaffenheit, so nennen wir den a priori zulässigen Punkt (C, e) des Phasenstreifens einen auch a posteriori zulässigen oder kurz fortsetzbaren Punkt und bezeichnen mit $\mathsf{M} = \mathsf{M}(C, e)$ die obere Grenze > 0 aller Zahlen $M(C, e)$, die mit Rücksicht auf die Definition des Massenprozentsatzes μ jedenfalls $\leqq 1$ sein muß. Gibt es aber kein $M = M(C, e) > 0$ von der erwähnten Beschaffenheit, so nennen wir den a priori zulässigen Punkt (C, e) des Phasenstreifens a posteriori nicht zulässig oder kurz nichtfortsetzbar und setzen $\mathsf{M}(C, e) = 0$. Endlich setzen wir $\mathsf{M}(C, e)$ auch dann $= 0$, wenn der Punkt (C, e) des Phasenstreifens nicht a priori zulässig ist. Damit ist in jedem der drei Phasenstreifen eine Funktion $\mathsf{M}(C, e)$ des Ortes definiert, die wir Massenschranke nennen wollen[1]).

Wir fügen nun zu den drei Phasenstreifen den Massenprozentsatz μ als dritte Koordinate eines kartesischen Bildraumes hinzu, so daß, sofern $\mathsf{M}(C, e) \neq 0$ ist, die periodischen Lösungen (93) durch die Punkte der auf dem Phasenstreifen senkrechten und diesen in dem Punkte (C, e) der Keplerschen Ebene $\mu = 0$ — in dem Fußpunkte von (60) — treffenden „μ-Strecke"

$$0 < \mu < \mathsf{M}(C, e) \qquad (C \text{ fest}, e \text{ fest}) \tag{98}$$

dargestellt werden. Auf einem festen „Lot" (vgl. S. 364) des Phasenstreifens $\mu = 0$ sind die in einer μ-Strecke enthaltenen periodischen Lösungen

durch $C =$ konst. charakterisiert (vgl. S. 373). Die andere Streifenkoordinate des Fußpunktes, nämlich e, heißt Eigenexzentrizität[26]) („ungestörte Exzentrizität") der μ-Strecke und die aus C und e gemäß (46) und (43) bestimmte Zahl (54) $[e = 0]$ bzw. (56) $[e \neq 0]$ die Eigenperiode („ungestörte Periode") der μ-Strecke; die Eigenperiode ist mit Rücksicht auf die erste und die sechste Forderung (S. 374) entweder gleich $T(+0)$ oder ein echter Teiler von $T(+0)$; es ist aber leicht von vornherein einzusehen, daß, sobald die Eigenexzentrizität gleich Null gewählt wird, der Grenzwert der Periode gleich der Eigenperiode sein muß, sofern die μ-Strecke überhaupt existiert. Die μ-Strecke heißt von der zweiten Sorte, wenn die Eigenexzentrizität $\neq 0$ ist, und sonst (d. h. wenn der Fußpunkt auf der Randgeraden $e = 0$ des Phasenstreifens $\mu = 0$ liegt) von der ersten Sorte. Die Gesamtheit derjenigen μ-Strecken, deren Fußpunkte auf der zu n gehörigen synchronen Kurve (91) des Phasenstreifens liegen, bezeichnen wir sinngemäß als die (zur Rationalzahl n gehörige) eigensynchrone Fläche, und die Gesamtheit aller μ-Strecken erster Sorte als die Randfläche der Sortenmannigfaltigkeit. Die Sortenmannigfaltigkeit besteht also einerseits aus der Folge der eigensynchronen Flächen und andererseits aus der Randfläche, die Folge der eigensynchronen Flächen aus den unterhalb der „Fläche" $\mu = \mathsf{M}(C, e)$ der Massenschranke liegenden Stücken der Folge der auf der Ebene $\mu = 0$ des Phasenstreifens senkrecht stehenden elliptischen Zylinderflächen (91) des dreidimensionalen Bildraumes (C, e, μ), endlich die Randfläche aus dem unterhalb der Massenschranke liegenden Teile der auf dem Phasenstreifen ebenfalls senkrechten Randebene $e = 0$, wobei μ durchweg als $\geqq 0$ vorausgesetzt ist. — Die Massenschranke $\mathsf{M}(C, e)$ ist auf der zu der festen Rationalzahl n gehörigen synchronen Kurve des Phasenstreifens eine Funktion $\mathsf{M}_n = \mathsf{M}_n(e)$ von e allein bzw. eine Funktion $\mathsf{M}'_n = \mathsf{M}'_n(C)$ von C allein, die man erhält, wenn man C bzw. e in $\mathsf{M}(C, e)$ mittels (91) als Funktion von e bzw. C ausdrückt. Die Verwendung von $\mathsf{M}_n(e)$ wird für uns bequemer sein als diejenige von $\mathsf{M}'_n(C)$ wäre. Es gilt freilich

$$\mathsf{M}_n(e) \equiv \mathsf{M}'_n(C) \tag{99}$$

vermöge (91), und zwar ist nach Definition

$$\mathsf{M}'_n(C) \equiv \mathsf{M}\left(C, \left|\sqrt{1 - \left(\frac{n - C\sqrt[3]{n}}{2}\right)^2}\right|\right) \qquad (n = \text{Rationalzahl}) \tag{100}$$

und

$$\mathsf{M}_n(e) \equiv \mathsf{M}\left(\frac{n + 2\,|\sqrt{1-e^2}|}{\sqrt[3]{n}}, e\right) \qquad (n = \text{Rationalzahl}). \tag{101}$$

[26]) Tisserand sagt excentricité propre. Das Wort Eigenexzentrizität ist zuerst wohl von Wilkens benutzt worden.

Auf der Randgeraden $e = 0$ des Phasenstreifens ist $\mathsf{M}(C, e)$ eine Funktion von C allein, die wir kurz mit $\mathsf{M}_0(C)$ bezeichnen wollen:

$$\mathsf{M}_0(C) \equiv \mathsf{M}(C, 0). \tag{102}$$

Mit Rücksicht auf die bei (91′) gemachte Bemerkung gilt für rationale n

$$\mathsf{M}_0\left(\frac{n+2}{\sqrt[3]{n}}\right) \equiv \mathsf{M}_n(0) \qquad (n = \text{Rationalzahl}), \tag{103}$$

doch ist die Funktion (102) nicht nur auf der überall dichten Punktfolge (91′), sondern auch in ihren Häufungsstellen, d. h. auf der ganzen Randgeraden des Phasenstreifens erklärt. Kennt man die Funktion (102) und die Funktionenfolge (101), so kennt man freilich die Funktion $\mathsf{M}(C, e)$ in allen Punkten der drei Phasenstreifen (durch Vorgabe von n und e ist schon eindeutig bestimmt, um welchen Phasenstreifen es sich handelt).

Ist μ^* irgendeine fest gewählte positive Zahl < 1, so bezeichnen wir die Gesamtheit aller in der zum Phasenstreifen parallelen Ebene $\mu = \mu^*$ gelegenen Punkte der vollen Sortenmannigfaltigkeit (d. h. der Gesamtheit aller μ-Strecken) — die also die bei dem festen Massenprozentsatz μ^* überhaupt existierenden Sortenlösungen des restringierten Dreikörperproblems darstellen und einerseits auf der Geraden $\mu = \mu^*$ der Randfläche (erste Sorte), andererseits auf der Folge von Ellipsenbogen liegen, in welchen die Folge der eigensynchronen Flächen durch die Ebene $\mu = \mu^*$ geschnitten wird (zweite Sorte) — kurz als den μ^*-Querschnitt der Sortenmannigfaltigkeit und sprechen beziehungsweise von dem μ^*-Querschnitt der Randfläche (erste Sorte) und von demjenigen der zu der Rationalzahl n gehörigen eigensynchronen Fläche (zweite Sorte). — Der μ^*-Querschnitt der ersten Sorte besteht nicht etwa aus der vollen Geraden $\mu = \mu^*$, $e = 0$ des dreidimensionalen Bildraums, sondern nur aus denjenigen ihrer Punkte, die unterhalb der Massenschrankenkurve

$$\mu = \mathsf{M}_0(C) \tag{104}$$

der Ebene $e = 0$ liegen, d. h. aus denjenigen Punkten der Geraden $\mu = \mu^*$, $e = 0$ des (C, e, μ)-Raumes, für welche

$$\mu^* < \mathsf{M}_0(C) \tag{105}$$

gilt, so daß der μ^*-Querschnitt der ersten Sorte bei noch so kleinem, aber festem μ^* aus mehreren Stücken bestehen, d. h. Lücken aufweisen muß, indem die Funktion $\mathsf{M}_0(C)$, wie wir sehen werden, nicht nur nicht $\equiv 1$ ist, sondern sogar Nullstellen hat; und Entsprechendes gilt [bereits wegen (103)] für die Folge der μ^*-Querschnitte der zweiten Sorte, sobald der Massenprozentsatz $\mu^* \neq 0$ beliebig fixiert wird.

In der Sortenproblematik handelt es sich nun nur um diejenigen periodischen Lösungen des restringierten Dreikörperproblems, die auf störungstheoretischem Wege erreichbar sind, d. h. um diejenigen Lösungen, die durch Punkte der zwischen der Keplerschen Ebene $\mu = 0$ und der Massenschranke $\mu = \mathsf{M}(C, e)$ liegenden Mannigfaltigkeit unseres dreidimensionalen Bildraumes geliefert sind. Dabei ist diese Störungsmannigfaltigkeit — genauer gibt es deren drei, entsprechend den drei Phasenstreifen — auf eine isoenergetische Weise erzeugt worden, indem die Jacobische Konstante C auf jeder μ-Strecke unabhängig von μ ist (so daß der sinngemäße Apparat durch das Maupertuissche Prinzip geliefert wird), während die Periode von μ abhängt. Wir werden später (S. 400) auch die analoge Störungsmannigfaltigkeit periodischer Lösungen betrachten, die auf eine isoperiodische Weise, d. h. derart definiert ist, daß auf einer μ-Strecke die Periode unabhängig von μ und die Jacobische Konstante entsprechend eine Funktion von μ wird; die zugehörige, d. h. die isoperiodische Massenschranke heiße $\mathsf{N}(C, e)$, zur Unterscheidung von der isoenergetischen Massenschranke $\mathsf{M}(C, e)$. Bis auf weiteres und sofern nichts weiter gesagt wird, also mit Ausnahme von § 10, handelt es sich stets um die isoenergetische Mannigfaltigkeit.

§ 5.

Die Massenschranke.

In diesem Paragraph wollen wir die Ergebnisse der Sortentheorie für spätere Zwecke (§§ 6 ff.) in unseren Bildraum (C, e, μ) einordnen. Was Existenzbeweise anlangt, kann sich so selbstverständlich nichts neues ergeben, es scheint daher, es handle sich dabei um ein bloßes Schematisieren. In Wirklichkeit erzwingt aber unser Modell eine ziemlich radikale Revision der Poincaréschen Problematik der Sorten [3]) [4]) [5]) und übrigens auch ihrer Poincaréschen Genealogie [5]), ja es ermöglicht überhaupt erst eine Sortengenealogie, weist ferner auf Fehlschlüsse, die seit gewissen vermeintlichen Existenzbeweisen von Poincaré und Schwarzschild (augenscheinlich unter dem Einfluß des bekannten Charlierschen Werkes) in der Literatur eine leider allzu große Verbreitung gefunden haben, von vornherein hin und leistet nicht nur bei dem Beweise der Nichtexistenz jener vermeintlich sichergestellten periodischen Lösungen gute Dienste, beleuchtet vielmehr einerseits auch das Lückenproblem der synodisch ganzzahligen Kommensurabilitäten in einem neuen Licht und erlaubt andererseits, über die bisher gar nicht beachteten Beziehungen des Lückenkomplexes und überhaupt der Sortenproblematik zu der Strömgrenschen Gruppenfortsetzung bindende Schlüsse zu ziehen (S. 389 ff.).

Die Kenntnis der Massenschranke $\mathsf{M}(C, e)$ verlangt nach S. 374 einerseits die Bestimmung der Massenschranke (104) erster Sorte und andererseits die Bestimmung der Folge

$$\mu = \mathsf{M}_n(e) \qquad (n = \text{Rationalzahl}) \tag{106}$$

der Massenschranken zweiter Sorte, so daß dieser Paragraph naturgemäß in zwei Teile (S. 378–382 und S. 382–385) zerfällt.

* * *

Die Vereinigung der Ränder $e = 0$ der drei Phasenstreifen $\mu = 0$, d. h. die zu $e = 0$ gehörige a priori zulässige Mannigfaltigkeit ist auf die Punkte des Kontinuums (59) oder — was wegen (65), (66) dasselbe ist — auf die Punkte des Kontinuums (68) eineindeutig bezogen, wobei die den Werten (60) entsprechenden Stellen (62) auszuschließen sind. Die zu m gehörige synodische primitive Periode ist dann durch (54) gegeben, und es handelt sich um eine synodische Kreisbewegung (um die Sonne $\xi = \eta = 0$) mit dem Radius a, der sich aus (86) berechnet, so daß also die fragliche, zu $\mu = 0$, $e = 0$ und zu gegebenem C gehörige, durch den Randpunkt $(C, 0)$ eines der drei Phasenstreifen dargestellte periodische Lösung offenbar [vgl. (70)]

$$\mu = 0,\ e = 0\colon\ \xi = \left(\sqrt[3]{\frac{m}{m+1}}\right)^2 \cos\frac{t}{m}, \quad \eta = \left(\sqrt[3]{\frac{m}{m+1}}\right)^2 \sin\frac{t}{m}; \quad C = \frac{3+\frac{1}{m}}{\sqrt[3]{1+\frac{1}{m}}} \tag{107}$$

ist. Da zu C nur im Falle $C < 3$ genau eine Bahn, im Falle $C > 3$ hingegen — entsprechend den drei Phasenstreifen — drei verschiedene Lösungen (107) gehören, während (107) durch m bei jedem m eindeutig bestimmt ist [indem C durch (70) im Reellen, d. h. gleichzeitig für alle drei Phasenstreifen $\mathfrak{S}_1^+$, $\mathfrak{S}_2^+$, $\mathfrak{S}^-$ uniformisiert wird], so wollen wir die Uniformisierungsvariable m auch in die Massenschranke (104) einführen und — unter Beachtung der Mehrdeutigkeit der Umkehrfunktion von (70) und nach Wahl des jeweilig richtigen Zweiges [vgl. (78)] — zur Abkürzung

$$\mathsf{M}^{(0)}(m) \equiv \mathsf{M}_0\left(\frac{3+\frac{1}{m}}{\sqrt[3]{1+\frac{1}{m}}}\right) \equiv \mathsf{M}_0(C) \tag{108}$$

setzen. Es handelt sich dann darum, diejenigen Werte von m zu bestimmen, bei welchen die zu dem Punkte

$$e = 0, \qquad C = \frac{3+\frac{1}{m}}{\sqrt[3]{1+\frac{1}{m}}} \tag{109}$$

der a priori zulässigen Mannigfaltigkeit gehörige Lösung (107) im Sinne von § 4 fortsetzbar, d. h. derart ist, daß es eine der auf S. 374 aufgezählten Bedingungen genügende Schar (93) von periodischen Lösungen gibt, die für $\mu \to 0$ in die Lösung (107) übergehen [die Massenschranke (108) wird mit Rücksicht auf die Definition von (104) für solche und nur für solche Werte von m positiv, d. h. $\neq 0$ sein].

Diese Störungsaufgabe kann heute außer mit Hilfe der Poincaréschen μ-Methode [27]) auch durch Benutzung des Apparats der Theorie der nichtlinearen Randwertprobleme [28]) sowie durch die Theorie der unendlich vielen Variablen [29]) gelöst werden. — Levi-Civita [30]) und Birkhoff [31]) haben auf Grund der Poincaréschen Methode bewiesen, daß die zu (107) gehörige Schar (93) jedenfalls dann existiert, wenn m nicht einen ausgeschlossenen Wert (62), aber auch nicht einen ganzzahligen Wert (87) hat, und genauer, daß es zu jedem, keinen der Werte (62), (87) enthaltenden m-Intervall

$$(110) \qquad m' \leqq m \leqq m'' \quad \left(s' < m' < m'' < s'', \text{ wenn } s', s'' = -\tfrac{1}{2}, 0, \pm 1, \ldots\right)$$

ein $\varepsilon > 0$ gibt derart, daß die Funktionen (93), (94) in dem Bereiche

$$(111) \qquad m' \leqq m \leqq m'', \quad 0 < \mu < \varepsilon, \quad -\infty < t < +\infty$$

vorhanden und regulär (und auch gleichmäßig stetig) sind. In unserer Sprache besagt dies u. a., daß, wenn m nicht einen der Werte (62), (87) hat, die Massenschranke (108) positiv, d. h. $\neq 0$ sein muß, und daß es darüber hinaus zu jedem, keinen der Werte (62), (87) enthaltenden m-Intervall (110) eine nur von m' und m'' abhängige Zahl $\varepsilon > 0$ mit

$$(112) \qquad \mathsf{M}^{(0)}(m) > \varepsilon > 0$$

in (110) gibt. Ob (108) im Falle (87) positiv oder aber $= 0$ ist, blieb dabei dahingestellt, da die Behandlung von (87) die Berechnung von höheren Termen verlangt, deren Berücksichtigung beim Beweise von (112) überflüssig ist. Herr Monroe Martin wird in einer demnächst in dem Amer. Journal of Math. erscheinenden Arbeit durch Berechnung der höheren Terme [32]) des näheren zeigen, daß für fast alle Zahlen der Folge (87) und dabei jedenfalls für

$$(113) \qquad m = j \qquad (j = 1, \pm 2, \pm 3, \pm 4)$$

---

[27]) a. a. O. [3]), chap. III.

[28]) L. Lichtenstein, Math. Zeitschr. **17** (1923), S. 62—110, ferner E. Hölder, Math. Zeitschr. **31** (1929), S. 197—257.

[29]) A. Wintner, Math. Annalen **96** (1926), S. 284—312; vgl. auch a. a. O. [1]), S. 15 ff., S. 50, ferner die weiter unten erwähnte Arbeit von Martin.

[30]) T. Levi-Civita, Annali di Math. (3) **5** (1901), S. 282—289.

[31]) a. a. O. [8]), § 11.

[32]) a. a. O. [1]), S. 17 bzw. S. 49 und insb. S. 51.

die Schar (93) nicht existiert, d. h. die periodische Störungsaufgabe unlösbar ist, was wir durch

$$\mathsf{M}^{(0)}(j) = 0 \tag{114}$$

ausdrücken können [ob es — jedenfalls nur endlich viele und von den Werten (113) jedenfalls verschiedene — Werte (87) gibt, für welche (114) nicht gilt, geht uns nicht an, da für uns nur die Werte (113) sowie große Werte von $|j|$ ein Interesse haben und j in den Anwendungen überhaupt nur einen der drei Werte $j = 1, 2, 3$ hat]. — Was nun endlich die vier Stellen (62) anlangt, so hat die Funktion (108) für diese Werte aus dynamischen Gründen keinen unmittelbaren Sinn, man kann aber fragen, wie sich (108) bei Annäherung an eine der Stellen (62) verhalten muß, wobei rechts- und linksseitige Annäherungen voneinander wohl zu unterscheiden sind, so daß es sich eigentlich nicht um vier, sondern um acht Stellen handelt:

$$m = +0, \quad m = -1-0, \quad m = -1+0, \quad m = -0 \tag{115}$$

und

$$m = +\infty, \quad m = -\infty, \quad m = -\tfrac{1}{2} - 0, \quad m = -\tfrac{1}{2} + 0, \tag{116}$$

entsprechend den Endpunkten der *vier* Gebiete (64) [vgl. Zeile III der Tabelle]. Die *drei* Stellen [vgl. (78)]

$$m = -\tfrac{1}{2}, \quad m = -\infty, \quad m = +\infty \tag{116}'$$

gehören, entsprechend den drei Phasenstreifen $\mathfrak{S}_1^+$, $\mathfrak{S}_2^+$, $\mathfrak{S}^-$, wegen (86) alle zu der Jupiterbahn $a = 1$ [und zwar $m = \pm\infty$ zu den Librationen (vgl. S. 368)] und über das Verhalten von (108) im Falle (116) ist gar nichts Sicheres bekannt. Anders steht es bei den vier Stellen (115), in welchen die vier Gebiete (64) ihren Anfang nehmen (vgl. Zeile XIV der Tabelle). — Was $m \to \pm 0$ anlangt, so ist dieser Grenzübergang wegen (86) mit $\alpha \to \pm 0$, also wegen $\alpha^2 = a > 0$ damit äquivalent, daß der synodisch direkt ($m > 0$) beziehungsweise synodisch retrograd ($m < 0$) durchlaufene Keplersche Kreis auf die Sonne zusammenschrumpft, und entsprechend ist $m \to -1 \pm 0$ wegen (86) mit $a \to +\infty$, also damit äquivalent, daß der Radius der siderisch direkt ($m < -1$) oder siderisch retrograd ($m > -1$) durchlaufenen Keplerschen Bahn ins Unendliche wächst. Wie groß nun dabei der Massenprozentsatz μ (< 1) auch sein mag, d. h. wie klein auch die Sonnenmasse $1 - \mu$ (> 0) ist, die Gravitationswirkung der Sonne wird gegenüber der Wirkung der Masse μ und der Zentrifugalkraft überwiegen, wenn man nur hinreichend nahe an die Sonne heranrückt, und entsprechend wird die Zentrifugalkraft in großen Entfernungen von den beiden Massen

überwiegend sein, so daß es vorauszusehen ist, daß bei noch so großem Massenprozentsatz μ bei hinreichend kleinem und bei hinreichend großem a, d. h. bei hinreichend kleinem $|m|$ und hinreichend kleinem $|m+1|$ gestörte kreisähnliche periodische Bahnen existieren müssen. In der Tat hat Birkhoff[34]) mittels eines elementaren Kunstgriffes, der der Regularisierungstheorie des restringierten Dreikörperproblems nachgebildet ist und freilich weniger weit als jene trägt, aber für diese Zwecke ausreicht, auf eine höchst einfache Weise gezeigt (genauer, es geht aus dem Birkhoffschen Beweise unmittelbar hervor), daß es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt derart, daß die Funktionen (93), (94) in dem Gebiete

$$(117) \qquad \left.\begin{matrix} 0 < m < \delta \\ -\delta < m < 0 \end{matrix}\right\}, \quad 0 < \mu < 1-\varepsilon, \quad -\infty < t < +\infty$$

vorhanden und regulär sind, und was die Umgebung von $m = -1$ anlangt, so folgt Entsprechendes genau so mittels der von Koopman[35]) für $\xi^2 + \eta^2 \to +\infty$ angegebenen Regularisierungstransformation, die sich übrigens für die vorliegenden Zwecke ebenfalls vereinfachen ließe. In unserer Sprache besagt dies u. a., daß, entsprechend den vier Hauptspalten der Tabelle, die Grenzgleichungen

$$(118) \quad M^{(0)}(+0) = 1, \quad M^{(0)}(-1-0) = 1, \quad M^{(0)}(-1+0), \quad M^{(0)}(-0) = 1$$

gelten, d. h. daß die Massenschranke in der Umgebung der in der Zeile XIV der Tabelle angegebenen Stellen so groß wird wie überhaupt möglich.

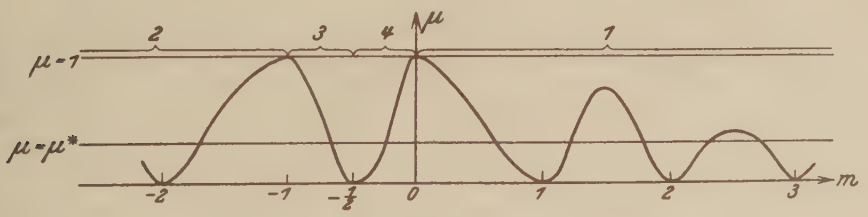


Fig. 1.

In Fig. 1 ist die Massenschranke $\mu = M^{(0)}(m)$ schematisch dargestellt; es ist dabei auch ein „μ^*-Querschnitt“ (vgl. S. 376) $\mu = \mu^*$ ausgezogen. Mehr als (112), (114), (118) ist heute nicht bekannt, die Kurve ist aber so gezeichnet, als ob ihre Stetigkeit sichergestellt wäre. In Fig. 2 ist dieselbe Massenschranke unter derselben Annahme als Funktion des Radius a dargestellt, wobei freilich wegen $\alpha = \pm\sqrt{a}$ zu jedem a zwei Funktions-

[34]) a. a. O. [8]), § 12.

[35]) B. O. Koopman, Trans. of Amer. Math. Soc. 29 (1927), S. 395—306.

werte gehören. Was $a = 1$, d. h. die drei Werte $(116)'$ anlangt, so ist die bei $(116)'$ gemachte Bemerkung nicht zu vergessen. In Fig. 3 ist endlich dieselbe Massenschranke als Funktion der Jacobischen Konstante dargestellt,

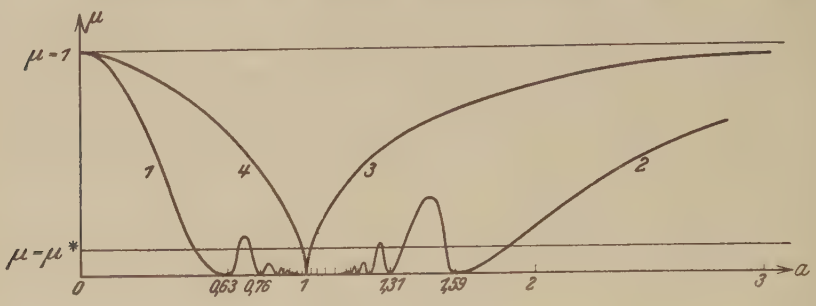


Fig. 2.

wobei (78), (89), (90) sowie die Zeilen VIII, IX der Tabelle zu beachten sind (während man zu Fig. 2 die Zeilen VI, VII vergleichen möge). — Fig. 3 stellt die Randfläche $e = 0$ (vgl. S. 375) der Sortenmannigfaltigkeit un-

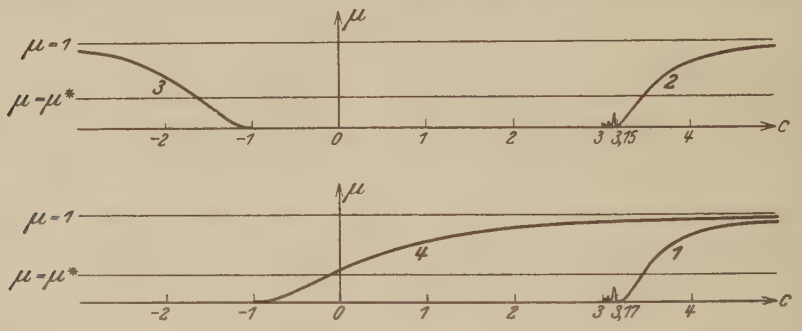


Fig. 3.

mittelbar dar, die zur Ordinatenachse parallelen, unterhalb der Massenschranke gelegenen Geradenstücke $C = \text{konst.}$ sind die „μ-Strecken“ (vgl. S. 374).

* * *

In den oben zitierten Untersuchungen über die erste Sorte ist der Jacobischen Konstante oft ein fester Wert zuerteilt und die Möglichkeit einer gleichmäßigen Aussage für ein Intervall von C-Werten öfters [vgl. z. B. bei [28])] nicht ausdrücklich formuliert (obwohl diese aus den Beweisen hervorgeht), was ganz im Rahmen der Sortenproblematik liegt, für die es sich doch nur um eine μ-Strecke handelt, die von einem *festen* Punkte der Ebene $\mu = 0$ ausgeht. Analog verhält es sich bei der zweiten Sorte,

zu der wir nun übergehen. Es wird sich dabei um die Untersuchungen von Birkhoff[36]) und Koopman[37]) handeln.

In dem durch $(n)_{e=0} > 1$ [vgl. (78) bzw. die Tabelle] charakterisierten Phasenstreifen

$$(119) \quad \mathfrak{S}_1^+: \qquad 3 < C < +\infty, \qquad 0 \leqq e < 1; \qquad (n)_{e=0} > 1,$$

der von seinem a priori zulässigen, in (119) überall dicht liegenden, aber nur aus der Folge der synchronen Kurven der Randgeraden $e = 0$ bestehenden Teile wohl zu unterscheiden ist, sind $\alpha = \alpha(C, e)$ und $n = n(C, e)$ eindeutige und stetige Funktionen des Ortes. Da C in $\mathfrak{S}_1^+$ stets > 3 ist, so existiert zu jedem C in $\mathfrak{S}_1^+$ das Zahlentripel (83), und es gilt dabei (vgl. S. 362)

$$(120) \qquad n(C, 0) = n_1(C) > 1,$$

d. h.

$$(121) \qquad 0 < \alpha(C, 0) = \alpha_1 = \alpha_1(C) < 1,$$

während $n_2(C)$ und $\alpha_2(C)$ zu $\mathfrak{S}_2^+$ und $n'(C)$ und $\alpha'(C)$ zu $\mathfrak{S}^-$ gehören [vgl. (78), (82) sowie die Tabelle]. Da $\alpha(C, e)$ in (119) stetig ist, so gibt es nach (121) zu jedem C ein hinreichend kleines $\varepsilon = \varepsilon_C > 0$ mit

$$(122) \qquad 0 < \alpha(C, e)^3 \cdot \left| \sqrt{1 - e^2} \right| < 1 \quad \text{für} \quad 0 < e < \varepsilon_C.$$

Dann ist aber die rechte Seite von (53) wegen $e < 1$ positiv, d. h. aus (122) und (53) folgt, wenn man auch noch (83) berücksichtigt,

$$(123) \qquad 0 < n_1(C) < n(C, e) \quad \text{für} \quad 0 < e < \varepsilon_C.$$

Birkhoff[36]) hat nun für $e \not\equiv 0$ bei festem C und unter der Annahme, daß es sich um „innere" ungestörte Bahnen (mit einer diese nach außen abgrenzenden Nullgeschwindigkeitskurve) handelt [den „äußeren" Fall behandelt Koopmann[37])], u. a. bewiesen, daß der Punkt (C, e) im Sinne von § 4 ein fortsetzbarer Punkt sein muß [d. h. daß $M(C, e) \neq 0$ gilt], wenn in dem Punkte (C, e) die Funktion $n(C, e)$ einerseits einen rationalen Wert n hat [d. h. wenn (C, e) überhaupt auf der für $e \neq 0$ wenigstens a priori zulässigen Mannigfaltigkeit, nämlich auf der zu der Rationalzahl n gehörigen synchronen Kurve liegt] und andererseits gewisse Ungleichungen

$$(124) \qquad \mathfrak{n}'(C) < n(C, e) < \mathfrak{n}_1(C)$$

befriedigt. Auf den Vergleich dieser Birkhoffschen Ungleichungen mit (84) und (123) bzw. mit den (123) in den übrigen Phasenstreifen entsprechenden Ungleichungen soll hier noch nicht eingegangen werden. Wir werden später untersuchen, ob die Birkhoffschen Ungleichheiten auch kleine Eigenexzentrizitäten zulassen.

---

[36]) a. a. O. [8]), § 20.

[37]) a. a. O. [35]), S. 323 ff.

Da das e-Lot $C = \text{konstant}$ nur von abzählbar vielen synchronen Kurven geschnitten wird, so kann sich freilich bei festem C nur eine diskrete [in der Ebene $C = \text{konst.}$ des Bildraumes (C, e, μ) bestenfalls überall dichte] Folge von solchen Birkhoffschen μ-Strecken (92) ergeben. Birkhoff zeigt nun u. a., daß bei festem C unendlich viele dieser μ-Strecken eine Länge haben, die größer ist als eine feste, von Null verschiedene Zahl. Der Nachweis dieses Satzes war das Hauptziel von Poincaré, als er in seiner letzten Arbeit[38]) den von Birkhoff[39]) später bewiesenen Fixpunktsatz postulierte; Birkhoff zeigt aber, daß man beim Beweis des Satzes den Fixpunktsatz entbehren und durch die weniger tief liegende Möglichkeit einer sehr einfachen Produktzerlegung[40]) der Ringtransformation[41]) ersetzen kann. Diese, zu einem festen C gehörigen Ergebnisse liegen in der Richtung der bekannten Poincaréschen, mit Rücksicht auf neuere Birkhoffsche Untersuchungen wahrscheinlich unrichtigen Vermutung, daß man, bei festem μ, jede Lösung des restringierten Dreikörperproblems für jedes feste t-Intervall durch eine Folge von zu demselben μ gehörigen periodischen Lösungen beliebig genau approximieren kann. Uns werden aber bei festem μ nicht diskrete Folgen von periodischen Lösungen, sondern Kontinua von solchen interessieren; zur Zeit interessiert uns also nicht so sehr die Folge der zu einem festen C gehörigen μ-Strecken, als das μ-Streckenkontinuum einer zu einer Rationalzahl n gehörigen eigensynchronen Fläche, in welcher C nicht einen festen Wert hat, sondern der Gleichung $n(C, e) = n$ genügt, so daß der Punkt (C, e) der a priori zulässigen Mannigfaltigkeit auf dem *Kontinuum* des zur Rationalzahl n gehörigen synchronen Ellipsenbogens und nicht auf einer *diskreten* Punktfolge [vgl. übrigens (58)] eines e-Lotes $C = \text{konst.}$ variiert. Nun muß zwar C bei der Ringtransformation[41]) (wie überhaupt im Rahmen der Maupertuisschen Sortenprobleme) ein für allemal fest gewählt werden, sie hängt aber von C jedenfalls stetig ab, so daß wir wieder (vgl. S. 382) Gleichmäßigkeitsaussagen machen können; es genügt, das über die Phasenstreifen (119) Gesagte mit den Birkhoffschen Beweisen zu vergleichen, um einzusehen, daß (106) — sofern kleine e überhaupt zulässig sind — bei jeder Rationalzahl $|n| > 1$ für hinreichend kleine Werte von e, etwa für

$$0 < e < \varepsilon^{(n)} \qquad (|n| = \text{Rationalzahl} > 1) \tag{125}$$

von Null verschieden ist:

$$M_n(e) > 0 \quad \text{für} \quad 0 < e < \varepsilon^{(n)} \qquad (|n| = \text{Rationalzahl} > 1), \tag{126}$$

[38]) H. Poincaré, Rend. Palermo **33** (1912), S. 375—407.
[39]) G. D. Birkhoff, Trans. of Amer. Math. Soc. **14** (1913), S. 14—22.
[40]) a. a. O. [8]), § 14.
[41]) a. a. O. [8]), § 13, § 15.

und daß es darüber hinaus zu jedem $\vartheta > 0$ und zu jedem n ein hinreichend kleines $\delta = \delta^{(n)}(\vartheta) > 0$ gibt derart, daß

(127) $\mathsf{M}_n(e) > \delta^{(n)}(\vartheta) > 0$ für $(0<)\,\vartheta < e < \varepsilon^{(n)}$ $(|n| = \text{Rationalzahl} > 1)$

gilt, was folgendes besagt: Die zu der Rationalzahl n gehörige Massenschranke, d. h. die durch die Gleichungen (106), (91) in Parameterdarstellung gegebene (eventuell nicht stetige) Raumkurve des Bildraumes (C, e, μ), die auf dem eigensynchronen elliptischen Zylinder (91) verläuft und die in diesem liegenden, d. h. zu der festen Rationalzahl n gehörigen, von der zu n gehörigen synchronen Kurve (91) des a priori $(\mu = 0)$ zulässigen Phasenraumes ausgehenden μ-Strecken (92) von oben begrenzt, muß oberhalb einer angebbaren Querebene $\mu = \delta^{(n)}(\vartheta) > 0$ [vgl. S. 379] verlaufen, solange diese Raumkurve von der auf dem Phasenstreifen $\mu = 0$ senkrecht stehenden Ebene $e = 0$ in einer noch so kleinen festen Entfernung $> \vartheta > 0$, aber von der Ebene $\varepsilon = 0$ nicht allzu weit entfernt, sondern in einer Entfernung bleibt, die kleiner ist als eine angebbare, nur von n (also nicht auch von ϑ) abhängige Zahl $\varepsilon^{(n)}(> \vartheta)$. — Ich bemerke noch, daß es auf der eigensynchronen Fläche (106) bei hinreichend kleinem fest gewähltem $\mu^* > 0$ ein hinreichend kleines Stück der Spur der Ebene $\mu = \mu^*$ auf der Fläche (106) gibt derart, daß die Funktionen (93), (94) in dem aus der unendlichen t-Achse und aus dem hinreichend kleinen Stücke der Spur bestehenden Bereiche gleichmäßig stetig und übrigens auch regulär sind, wie es aus den Beweisen unmittelbar hervorgeht.

§ 6.

Das Prinzip der Massenfixierung.

In diesem Paragraphen und auch in den beiden folgenden beschränken wir uns der Durchsichtigkeit halber (und nur aus diesem Grunde) auf den Phasenstreifen $\mathfrak{S}_1^+$; die beiden anderen Phasenstreifen ($\mathfrak{S}_2^+$ und $\mathfrak{S}^-$) sollen erst in § 9 besprochen werden.

Es bezeichne $\mathfrak{F}$ die Gesamtheit aller Sortenlösungen, die für $\mu \to 0$ in durch Punkte von (119) dargestellte ungestörte $(\mu = 0)$ periodische Lösungen übergehen, d. h. die Gesamtheit aller μ-Strecken, deren Fußpunkt (vgl. S. 374) in $\mathfrak{S}_1^+$ liegt. $\mathfrak{F}$ besteht einerseits aus der Mannigfaltigkeit der μ-Strecken erster Sorte, d. h. aus der Randfläche (vgl. S. 375), die $\mathfrak{W}$ heißen möge, andererseits aus der Mannigfaltigkeit der μ-Strecken zweiter Sorte, d. h. aus der Folge der zu Rationalzahlen $n > 1$ gehörigen eigensynchronen Flächen, die durch $\mathfrak{Z}$ bezeichnet werden soll, so daß $\mathfrak{F} = \mathfrak{W} + \mathfrak{Z}$ gilt. $\mathfrak{W}$ ist also das durch die beiden Geraden $\mu = 0$, $C = 3$ und die Massenschrankenkurve begrenzte Stück der in Fig. 3 dargestellten Ebene $e = 0$ und zerfällt, wie aus Fig. 3 und bequemer aus Fig. 1 oder Fig. 2

ersichtlich ist, wegen (114), (112) in unendlich viele, voneinander durch die Punkte (89) der C-Achse getrennte „Wellen", die $\mathfrak{W}_0, \mathfrak{W}_1, \mathfrak{W}_2, \ldots$ heißen mögen. $\mathfrak{W}_j$ liegt zwischen $(C)_{m=j-1}$ und $(C)_{m=j}$ [vgl. (89)], so daß also $\mathfrak{W}_0$, das von den massennahen Bahnen ausgeht und bis zu der Hekubakommensurabilität $m=1$ reicht, zwischen $(C)_{m=0}=+\infty$ und $(C)_{m=1}=3{,}17$ liegt, während alle übrigen $\mathfrak{W}_j$ in dem Rechteck von $C=(C)_{m=1}=3{,}17$ bis $C=(C)_{m=+\infty}=3$ Platz finden (vgl. Fig. 3). Da sehr kleine m Mondbahnen entsprechen und $m=1$ und $m=2$ die Kommensurabilitäten des Hekuba- bzw. Hildatypus darstellen, so heiße $\mathfrak{W}_0$ Mondwelle, $\mathfrak{W}_1$ Hekubawelle und $\mathfrak{W}_2$ Hildawelle. — Bezeichnet man noch entsprechend die zur Rationalzahl $n>1$ gehörige eigensynchrone Fläche [d. h. das durch den Ellipsenbogen (91) der (C,e)-Ebene $\mu=0$ und durch die Raumkurve (91) (106) des (C,e,μ)-Raumes (vgl. S. 374 ff.) begrenzte Stück der auf der Ebene $\mu=0$ senkrechten elliptischen Zylinderfläche (91) des (C',e,μ)-Raumes] — sofern sie überhaupt existiert — mit $\mathfrak{Z}_n$, so haben wir zusammenfassend

$$(128)\qquad \mathfrak{F}=\mathfrak{W}+\mathfrak{Z},\quad \mathfrak{W}=\sum_{j=0}^{\infty}\mathfrak{W}_j,\quad \mathfrak{Z}=\sum_{n}\mathfrak{Z}_n\quad (n=\text{Rationalzahl}>1).$$

Die auf S. 376 „μ^*-Querschnitte" genannten Geraden- bzw. Ellipsenbogenstücke, die die Spur der Ebene $\mu=\mu^*$ auf den Flächenstücken $\mathfrak{W}_j$ bzw. $\mathfrak{Z}_n$ darstellen, sollen mit $\mathfrak{W}_j(\mu^*)$ bzw. $\mathfrak{Z}_n(\mu^*)$ bezeichnet werden, so daß

$$(129)\qquad \mathfrak{F}(\mu^*)=\sum_{j=0}^{\infty}\mathfrak{W}_j(\mu^*)+\sum_{n}\mathfrak{Z}_n(\mu^*)\qquad (1<n=\text{Rationalzahl})$$

wegen (128) die Gesamtheit aller bei dem festen Werte μ^* des Massenprozentsatzes in unserer Sortenmannigfaltigkeit liegenden periodischen Lösungen darstellt. $\mathfrak{W}_j(\mu^*)$ ist das in der Welle $\mathfrak{W}_j$ liegende Stück der Geraden $\mu=\mu^*$ der Randfläche (vgl. die Figuren) — wenn es ein solches Stück überhaupt gibt [und sonst die leere Menge].

Da $\mathfrak{W}_j(\mu^*)$ in der Ebene $e=0$ verläuft, in welcher $\mathfrak{Z}_n(\mu^*)$ höchstens *einen* Punkt hat, so können $\mathfrak{W}_j(\mu^*)$ und $\mathfrak{Z}_n(\mu^*)$ nur einen einzigen Punkt oder gar keinen gemein haben, und gibt es einen gemeinsamen Punkt, so ist dieser wegen (91′) offenbar der Punkt

$$(130)\qquad C=\frac{n+2}{\sqrt[3]{n}},\qquad e=0,\qquad \mu=\mu^*.$$

Haben $\mathfrak{W}_j(\mu^*)$ und $\mathfrak{Z}_n(\mu^*)$ den einzig möglichen gemeinsamen Punkt (130) tatsächlich gemein und sind die Funktionen (94) nicht nur in einer hinreichend nahen Umgebung von (130) auf $\mathfrak{W}_j(\mu^*)$ und auch auf $\mathfrak{Z}_n(\mu^*)$ [vgl. (111) und die Schlußbemerkung von § 5], sondern auch in dem Punkte (130) selbst, also in einer hinreichend kleinen, (130) enthaltenden offenen Umgebung des Kurvenzuges $\mathfrak{W}_j(\mu^*)+\mathfrak{Z}_n(\mu^*)$ analytische Funk-

tionen des Ortes auf $\mathfrak{W}_j(\mu^*) + \mathfrak{Z}_n(\mu^*)$, so wollen wir kurz sagen, $\mathfrak{Z}_n(\mu^*)$ sei die analytische Fortsetzung von $\mathfrak{W}_j(\mu^*)$. — Es sei hier erwähnt, daß es ganz sinnlos wäre, von einer analytischen Fortsetzung in $\mathfrak{F}$ nicht nur bei einem festen $\mu = \mu^*$, sondern auch bei variablem Massenprozentsatz zu reden, da doch sonst die *ganze* Sortenmannigfaltigkeit von vornherein als eine einzige analytische Mannigfaltigkeit zu gelten hätte, indem ja alle μ-Strecken von der Ebene $\mu = 0$ ausgehen; genauer, es handelt sich um analytische Funktionen dreier Veränderlichen (C, e, μ), die nur auf zweidimensionalen Flächenstücken erklärt sind. Außerdem ist ja μ keinesfalls verfügbar, sondern ein Datum des Problems (z. B. bei den kleinen Planeten im Falle Jupiter $= \frac{1}{1048 \cdot 35}$), das die Differentialgleichungen überhaupt festlegt, also nicht eine Integrationskonstante ist[42]).

Zunächst ist es nun klar, daß (wie auch $\mu^* \neq 0$ gewählt sein mag) kein $\mathfrak{W}_j(\mu^*)$ die analytische Fortsetzung von $\mathfrak{Z}_n(\mu^*)$ sein kann, wenn der Zeiger von $\mathfrak{Z}_n(\mu^*)$ dem Hillschen Periodenquotienten $m = \frac{1}{n-1}$ einen ganzzahligen, d. h. kritischen Wert erteilt (vgl. die erste Spalte der Tabelle). In der Tat können $\mathfrak{W}_j(\mu^*)$ und $\mathfrak{Z}_n(\mu^*)$ bei ganzzahligem m und noch so kleinem μ^* mit Rücksicht auf (114) [vgl. auch (103), (108)] nicht einmal einen gemeinsamen Punkt haben. Im Falle eines Lückenzentrums ist also die erste Sorte bei keinem $\mu^* \neq 0$ die analytische Fortsetzung der ersten Sorte. — Dies gilt übrigens wahrscheinlich auch dann, wenn m kein Lückenzentrum, d. h. nicht ganzzahlig ist. Für hinreichend kleine μ^* (d. h. diesmal nicht für alle μ^*) ist dies trivial. Es sei nämlich die Rationalzahl n gegeben und $m = \frac{1}{n-1}$ nicht ganzzahlig. $\mathfrak{Z}_n(\mu^*)$ kann mit Rücksicht auf (130) höchstens mit demjenigen $\mathfrak{W}_j(\mu^*)$ einen Punkt gemein haben, für welches $j - 1 < m < j$ gilt. Nach (111) kann man ferner zu jedem $\delta > 0$ ein $\varepsilon > 0$ finden derart, daß die Funktionen (94) in dem Gebiete $j - 1 + \delta < m < j - \delta$, $\mu = \mu^*$ der Randebene $e = 0$ regulär sind, wenn $\mu^* < \varepsilon$ gewählt wird. Ist also $\mu^* < \varepsilon$ (die Existenz der Sortenlösungen ist überhaupt stets nur für hinreichend kleine μ bewiesen! vgl. S. 379, S. 384), so kann $\mathfrak{Z}_n(\mu^*)$ unmöglich die analytische Fortsetzung von $\mathfrak{W}_j(\mu^*)$ sein, da doch sonst der gemeinsame Punkt (130) [sofern dieser überhaupt existiert] nicht ein regulärer, sondern ein Verzweigungspunkt mindestens einer der Funktionen (94) wäre. — Man möge in diesem Zusammenhang auch die diesbezüglichen Ausführungen der ersten Mitteilung (über Verzweigungen im allgemeinen) vergleichen.

* * *

[42]) Vgl. a. a. O. [1]), S. 31 ff.

Es soll nun zunächst durch Widerlegung der Poincaréschen Genealogie[5]) und dann in den folgenden Paragraphen durch weitergehende Anwendungen gezeigt werden, was durch die oben zugrunde gelegte methodische Wendung, wonach man die Sortenmannigfaltigkeit am besten nicht als ein einziges analytisches Gebilde, sondern nur als Vereinigung ihrer zu den verschiedenen, aber festen μ^* gehörigen μ^*-Querschnitte auffaßt[42]) — ich will das kurz „Prinzip der Massenfixierung“ nennen —, sowohl mathematisch als auch astronomisch gewonnen ist.

Es sei hier nicht im einzelnen wiederholt, was ich auf S. 29 ff. der vorläufigen Mitteilung[1]) über die Zirkelschlüsse der Poincaréschen Theorie der kritischen Kommensurabilitäten gesagt habe. Diese Fehlschlüsse rühren ausschließlich davon her, daß bei Poincaré der Begriff der Massenschranke fehlt. Es handelt sich darum, daß Poincaré[5]) bei der ersten Sorte an Stelle einer Massenschranke mit dem völlig vagen Begriffe eines „hinreichend kleinen μ“ operiert und daher nicht bemerkt, daß es dabei von C abhängt, *was* für μ „hinreichend klein“ ist. Poincaré nimmt dadurch die Fourierreihen (93) auch oberhalb der Massenschranke in Anspruch, wo sie doch nicht einmal im Sinne einer analytischen Fortsetzung existieren, geschweige denn konvergieren können, und hält damit das, was er so eben beweisen will, daß nämlich die zweite Sorte, wenigstens für hinreichend kleine Werte des Massenprozentsatzes, die analytische Fortsetzung der ersten Sorte ist, also den Hauptsatz seiner Genealogie, implizite von vornherein für selbstverständlich (des näheren vgl. S. 30 der vorläufigen Mitteilung[1])). Loc. cit. (vgl. auch S. 48) konnte ich nur zeigen, daß der Poincarésche Hauptsatz seiner Sortengenealogie auf solchen Zirkelschlüssen beruht.

Oben (S. 387) ist nun aber die Poincarésche Sortengenealogie endgültig widerlegt worden. Da die Sache ganz trivial ist, so hätte das freilich auch Poincaré bemerken können, wenn auch er mit einem dreidimensionalen Modell operiert und dessen dritte Dimension nicht dadurch ausgeschaltet hätte, daß er das ganze Problem sozusagen in die nur zweidimensionale „Ebene“

$$\mu = \text{„hinreichend klein“} \tag{131}$$

unseres (C, e, μ)-Raumes abplatten läßt[5]). In Wirklichkeit steht aber, wie wir in § 7 sehen werden, nicht nur die Verwendung der inhaltslosen Begriffsbildung (131), sondern bereits die prinzipielle Einstellung der in den Méth. Nouv. dargestellten ganzen Poincaréschen Problematik im Hintergrund.

* * *

Was die Massenschranke anlangt, ist folgendes zu beachten. Aus Fig. 2 oder Fig. 1 geht hervor, daß die Lückenzentra $m = j$, die für $\mu = 0$ nur noch *Punkte* sind, beim Übergang von $\mu = 0$ zu $\mu = \mu^*$ (wie klein dabei auch $\mu^* > 0$ sein mag) sich dank des Prinzips der Massenfixierung zu

oberhalb der Massenschranke liegenden Lücken*gebieten* ausbreiten, sofern es feststeht, daß die Massenschranke für $m = j$ stetig ist, d. h. daß außer (114) auch $\mathsf{M}^{(0)}(j \pm 0) = 0$ gilt (was wir nicht bewiesen haben), und Entsprechendes gilt auch bei Annäherung an die Jupiterbahn $a = 1$ [vgl. Fig. 2, ferner die Zeilen IV—VIII der Tabelle auf S. 369, endlich das auf S. 380 über die Librationen Gesagte]. Ganz analog liegen die Verhältnisse bei der zweiten Sorte.

Sieht man also von der Frage nach der dynamischen Zulässigkeit des Restringierens, d. h. von der Vernachlässigung der Jupiterexzentrizität, der Einfachheit halber[43]) ab, so gewinnt man den Eindruck, daß die konsequente Zugrundelegung der Massenschranke und des Prinzips der Massenfixierung das astronomische Problem wenigstens der synodisch ganzzahligen Lückengebiete der mittleren Bewegungen der kleinen Planeten, ganz im Gegensatz zu (131), quantitativ und rein analytisch zu fassen vermag[44]) [da doch die Jupitermasse nicht „hinreichend klein", d. h. nicht verfügbar, sondern gleich $\frac{1}{1047 \cdot 35}$ ist, im Einklang damit, daß die astronomischen Lückengebiete nicht exakte Kommensurabilitäten, sondern ausgebreitete Zonen[45]) sind]. — In Wirklichkeit ist aber in dieser Richtung nichts zu erwarten, und zwar deshalb nicht, weil der Poincarésche Sortenbegriff eine äußerst willkürliche mathematische Konstruktion ist und daher auch seine Massenschranke, wie wir gleich sehen werden, keine reale genealogische Bedeutung hat. — [Übrigens wäre es für manche Zwecke bequem, die μ-Strecken nur so lange zu konstruieren, d. h. die sortenmäßige analytische Fortsetzung nur so lange zu treiben, als die Syzygienachse orthogonal passiert wird, und die Massenschranke in diesem Sinne abzuändern, also zu verkleinern; doch soll darauf hier nicht näher eingegangen werden, da dadurch für die vorliegenden Zwecke gar nichts gewonnen werden könnte.]

§ 7.

Die Gruppe g und der Hekubakomplex.

Damit sind wir am springenden Punkt angelangt.

Der Poincarésche Gedankengang beim Beweise des (vorhin widerlegten) Hauptsatzes, wonach die zweite Sorte die analytische Fortsetzung der ersten

[43]) Vgl. aber die unter [6]) an zweiter Stelle erwähnte Arbeit von Wilkens (1913).

[44]) Klose [6]) behandelt die nicht kritischen Kommensurabilitätslücken.

[45]) Zur Orientierung vgl. G. Stracke, Ergebnisse der exakten Naturwissenschaften, 4 (1925), S. 1ff. — Da es für die verschiedenen hochzahlig kritischen Kommensurabilitäten keine *verschiedenen* Lückenzonen gibt, sondern letztere zu einem einzigen bis an die Librationszone ($a = 1$) der Trojaner reichenden planetenfreien Ring zusammenschmelzen, so liegt die Vermutung nahe, daß bereits die Massenschranke erster Sorte sich für $a = 1$ stetig verhält, d. h. für $a \to 1$ gegen Null konvergiert, wie es in Fig. 2 angegeben, aber z. Z. nicht bewiesen ist (vgl. S. 381 oben). — Vgl. aber S. 392 weiter unten.

ist[5]), ist an sich sehr einfach[46]): Da die Fourierreihen der ersten Sorte in der Umgebung der Lückenzentra ihre Gültigkeit verlieren und in Reihen mit nicht verschwindender Eigenexzentrizität übergehen und da sie dadurch (man vergleiche jedoch S. 388) dennoch die Existenz von periodischen Lösungen „anzeigen", so müssen letztere periodische Lösungen von der zweiten Sorte sein, da doch die Sortenmannigfaltigkeit nur periodische Lösungen der ersten und der zweiten Sorte enthält, und die periodischen Lösungen in der Nähe der Lückenzentra unmöglich von der ersten Sorte sein können. Folglich ist die zweite Sorte die analytische Fortsetzung der ersten. — Da aber dieses Schlußresultat nach § 6 gewiß nicht zutreffend sein kann, so haben wir die folgende Alternative: Entweder sind die Fourierreihen in dem von Poincaré betrachteten kritischen Gebiete[46]) gar nicht vorhanden, oder es muß ein Versehen von prinzipieller Bedeutung vorliegen. Ich werde zeigen, daß jedenfalls ein Versehen grundsätzlicher Art vorliegt.

Ich wiederhole, daß das Ergebnis bei Poincaré auf eine recht spontane Weise herauskommt: *weil* die fraglichen periodischen Lösungen (sofern sie freilich existieren) *nicht* von der ersten Sorte sind, *so müssen* sie von der zweiten Sorte sein, da es doch (sofern sie existieren) eine dritte Möglichkeit gar nicht gibt. Diese Schlußweise ist freilich die Grundlage auch derjenigen Variante des Beweises des (wie wir wissen, nicht haltbaren) Hauptsatzes, bei welchem Poincaré die Delanuaysche Methode benutzt[47]). Ist aber diese Schlußweise überhaupt zulässig? Gibt es vielleicht nicht auch periodische Lösungen, die gar nicht in der Sortenmannigfaltigkeit liegen, d. h. die auf störungstheoretischem Wege, durch Übergang von $\mu = 0$ zu $\mu \neq 0$, d. h. durch Konstruktion eines μ-Lotes gar nicht erreichbar sind? Gibt es solche Lösungen, so ist die Poincarésche Problematik, wie sie in den Méth. Nouv. dargestellt ist, nicht erschöpfend. Die Librationslösungen würden dies noch nicht überzeugend in Evidenz setzen, da man die Librationen an und für sich noch zu den Sortenlösungen zählen könnte; vgl. S. 368. Hingegen kann man durch Vergleich der Strömgrenschen Gruppe g mit dem Hekubakomplex zum Ziele gelangen, was ich in der vorläufigen Mitteilung[1]) nur vermutet habe (S. 46), jetzt aber beweisen werde.

* * *

Die Mondwelle $\mathfrak{W}_0$, die in Fig. 3 von $C = +\infty$ bis $C = 3{,}17$ reicht, enthält bei noch so großem, aber festem Massenprozentsatz μ (< 1) ein Stück der von der μ-Achse in der Entfernung μ verlaufenden Geraden,

---

[46]) a. a. O. [5]), S. 179 ff., insb. S. 186.

[47]) a. a. O. [5]), S. 186–188; vgl. auch a. a. O. [4]), S. 355 ff.

da doch die Höhe von $\mathfrak{W}_0$ für $C \to +\infty$ mit Rücksicht auf die erste der vier Gleichungen (118) gegen 1 ($> \mu$) konvergiert. D. h., der μ-Querschnitt $\mathfrak{W}_0(\mu)$ von $\mathfrak{W}_0$ ist bei jedem festen μ für hinreichend große C jedenfalls vorhanden. Der Scharparameter der durch die Punkte von $\mathfrak{W}_0(\mu)$ repräsentierten periodischen Lösungen ist die Jacobische Konstante C, und die Schar ist, da μ fest gewählt wurde, einparametrig, d. h. es gilt für hinreichend große C

$$(131) \qquad \xi = \sum_{k=-\infty}^{+\infty} \alpha_k^{(\mu)} \exp \frac{2\pi t i k}{T^{(\mu)}}, \qquad \eta = \sum_{k=-\infty}^{+\infty} \beta_k^{(\mu)} \exp \frac{2\pi t i k}{T^{(\mu)}},$$

wobei die Funktionen

$$(132) \qquad \bar{\alpha}_{-k}^{(\mu)} = \alpha_k^{(\mu)} = \alpha_k^{(\mu)}(C), \quad \bar{\beta}_{-k}^{(\mu)} = \beta_k^{(\mu)} = \beta_k^{(\mu)}(C), \quad T^{(\mu)} = T^{(\mu)}(C) > 0$$

bei festem μ nur von C abhängen und für hinreichend große C mit Rücksicht auf (117) gewiß regulär sind. Man kann die Existenz dieser Schar von periodischen Lösungen (bei festem μ für hinreichend große C) auch ohne Benutzung des Anschmiegungsprinzips (vgl. S. 379) mittels der direkten Methode der unendlich vielen Variablen, d. h. nach der Methode der unbestimmten Koeffizienten sicherstellen, wie ich es in dem Hillschen Grenzfall getan habe[48]); die ersten Glieder von (131) [wobei C bei Annäherung an die Masse nicht gegen einen endlichen Grenzwert, sondern gegen $+\infty$ konvergiert] sind so von Fischer-Petersen[49]) angesetzt worden, und meine Methode ergibt dabei den direkten Existenzbeweis. So folgt, daß die massennahe Schar direkter periodischer Lösungen bei jedem μ eindeutig bestimmt ist und daß die Differenz $\alpha_1^{(\mu)} - \beta_1^{(\mu)}$ sowie die Koeffizienten $\alpha_2^{(\mu)}$, $\beta_2^{(\mu)}$, $\alpha_3^{(\mu)}$, $\beta_3^{(\mu)}$, ... Potenzen von $C^{-\frac{1}{2}}$ enthalten, durch welche $\alpha_1^{(\mu)}$ und $\beta_1^{(\mu)}$ noch nicht teilbar sind, so daß die Bahn (131) bei noch so großem, aber festem μ um so mehr kreisähnlich wird, je größer C gewählt ist (dem entspricht es, daß zu $C \to +\infty$ unbegrenzte Annäherung an die Masse gehört; vgl. S. 382). — $\mathfrak{W}_0(\mu)$ ist also in der Terminologie der ersten Mitteilung das Stück einer (zu μ gehörigen) Gruppe, und zwar ihr „astronomisch selbstverständliches" Anfangsstück. Die fragliche Gruppe ist offenbar die Kopenhagener Gruppe g,[20]) wobei ich die Strömgrensche Gruppenbezeichnung[18]), die sich eigentlich nur auf den — übrigens gewiß ausschlaggebenden — Fall $\mu = \frac{1}{2}$ bezieht, freilich bei

---

[48]) Math. Zeitschr. 24 (1925), S. 259–265, und darüber hinausgehend 30 (1930), S. 211–227.

[49]) a. a. O. [22]), S. 389ff. Hill hat nur eine reziproke Entfernung, die er also mittels des Jacobischen Integrals eliminieren kann. Doch ist dieser Umstand unwesentlich, da die Thielesche Transformation beide reziproken Distanzen des eigentlichen restringierten Problems ohne weiteres behebt.

jedem μ in Anspruch nehme. — Über die weitere, d. h. *nach* dem astronomisch selbstverständlichen Gruppenstück erfolgende Entwicklung der Gruppe g (wobei μ beliebig, aber fest ist) gibt das Strömgrensche Abschlußprinzip[10]) Aufschluß, das ich in der ersten Mitteilung[2]) mathematisch legalisiert habe, so daß die numerischen Ergebnisse als mathematisch fundiert gelten können.

Die Strömgrensche Gruppe g enthält nun Bahnen auch mit $C < (C)_{m=1} = 3{,}17$ [vgl. (24) mit (133)], während die Massenschranke in dem Lückenzentrum $m = 1$ gewiß verschwindet. Folglich ist die „horizontale" Gruppenfortsetzung auch über die Hekubalücke hindurch möglich, die demnach bei gruppenmäßiger Fortsetzung keine reale genealogische Bedeutung hat. Damit ist nicht nur gezeigt, daß nicht alle periodischen Lösungen in der Sortenmannigfaltigkeit gelegen, d. h. störungstheoretisch, m. a. W. durch „vertikale" analytische Fortsetzung *erreichbar* sind, es folgt vielmehr auch die innere Inkonsistenz der Sortenproblematik, da die Sortentheorie durch sortenmäßige (d. h. vertikale) analytische Fortsetzung das astronomisch selbstverständliche Stück der Gruppe g zu konstruieren vermag, das nach gruppenmäßiger (d. h. horizontaler) analytischer Fortsetzung aus der Sortenmannigfaltigkeit hinausragt[50]) [dabei haben wir nur (114) für $j = 1$ benutzt und die Stetigkeit der Massenschranke (vgl. S. 381) keinesfalls vorausgesetzt]. Die Sortenmannigfaltigkeit ist also nicht einmal etwas Abgeschlossenes, sondern ein recht gekünstelt aufgebautes Gebilde[50]), indem die Gruppe g zunächst unterhalb, dann aber oberhalb der Massenschranke verläuft. Man könnte zunächst allerdings meinen, daß die Gruppe nach Passieren der Mondwelle und der Hekubalücke in die Hekubawelle hineindringt und dadurch wieder unterhalb der Massenschranke zu liegen kommt. Man beachte aber, daß *alle* von der Mondwelle verschiedenen Wellen zwischen $C = 3{,}17$ und $C = 3$ liegen (vgl. S. 372), während die Gruppe g Bahnen auch mit $C < 3$ enthält [vgl. (24) mit VIII ff. auf S. 393]. Daß die Gruppe nach Passieren der Mondwelle und im Prinzip vielleicht schon früher jeden Kontakt mit der Sortenproblematik verliert, tritt aber auch auf eine viel überraschendere Weise in Evidenz. Dazu wollen wir nun übergehen.

§ 8.

Bestätigung der Theorie durch Reduktion der Darwinschen Satellitenfamilien auf die Kopenhagener Gruppe g.

Die Werte der Kopenhagener Jacobischen Konstante K für die ersten fünfzehn Bahnen der Gruppe g sind die folgenden[51]):

---

[50]) Funktionentheoretisch vgl. hierzu a. a. O. [1]), S. 32—33.

[51]) a. a. O. [12]), S. 37—40 oder a. a. O. [18]), S. 15—21.

(133)
$$K: \begin{cases} \text{I} & \text{II} & \text{III} & \text{IV} & \text{V} & \text{VI}^* \\ 200 & 64 & 24 & 14{,}95 & 15{,}35 & 13{,}94 \\ \text{VII} & \text{VIII} & \text{IX} & \text{X}^* & & \\ 13{,}19 & 11{,}21 & 10{,}04 & 10{,}63 & & \\ \text{XI} & \text{XII} & \text{XIII} & \text{XIV} & \text{XV}^* & \\ 10{,}24 & 11{,}22 & 1426 & 13{,}60 & 10{,}98. & \end{cases}$$

Es sind auch weitere Bahnen der Gruppen berechnet[52]), die zum Teil bereits veröffentlicht sind[51]). Ich habe nur zwei Dezimalen angegeben, da die Darwinschen Werte nicht einmal so genau sind. In (133) sind die Bahnen in der Reihenfolge angegeben, in der sie analytische Fortsetzungen voneinander sind. (133) enthält, entsprechend den drei Zeilen in (133), die ersten drei Phasen der Entwicklungsgeschichte der Gruppe g, d. h. die drei mit Stern versehenen Bahnen VI*, X*, XV* (und vor XV* auch nur diese Bahnen) sind Stoßbahnen, die den Umlaufssinn umkehren. Doch ist damit der qualitative Entwicklungsgang der Gruppe nur sehr unvollständig charakterisiert, da die Schleifen, die auch nach Spitzen von Nichtstoßbahnen entstehen, sowie Selbstdurchsetzungen eine sehr große Rolle spielen, ebenso wie umgekehrt das Verschwinden der Schleifen bzw. Selbstdurchsetzungen in der „v. Haerdtlschen ($K = 13{,}67$) Gegend“ XIII bis XIV. Ich muß also auf die Strömgrenschen Figuren verweisen[53]). — Die erste Phase der Entwicklung (I bis VI*) ist nichts weiter als die Darwinsche Satellitenfamilie A[54]), nur daß μ bei Darwin nicht $= \frac{1}{2}$, sondern $= \frac{10}{11}$ ist [und offenbar nicht $\frac{1}{11}$, vgl. S. 357 oben; bei den Darwinschen planets A[55]) handelt es sich dementsprechend gar nicht um die Gruppe g, sondern um die Gruppe i, was nur im Falle $\mu = \frac{1}{2}$ keinen eigentlichen Unterschied ausmachen würde]. Strömgren[65]) hatte nun die grundlegende Bemerkung gemacht, daß die Darwinschen Satellitenfamilien B[56]), C[57]) ebenfalls der Gruppe g angehören. Die Sache steht nämlich so. In Darwins großer Arbeit ist die Stoßbahn der ersten Phase (m. a. W. die erste Stoßbahn der Gruppe g) übersehen worden[58]), die verschiedene Unstimmigkeiten

---

[52]) Die Bahnen I bis VI sind auch in allen ihren Einzelheiten veröffentlicht worden, s. C. Burrau und E. Strömgren, Astr. Nachr. 200 (1915), S. 313—330 = Publ. Købn. Observ. Nr. 21.

[53]) a. a. O. [52]) oder a. a. O. [11]), S. 979 ff.

[54]) a. a. O. [16]), S. 69 und plate IV.

[55]) a. a. O. [16]), S. 70 und plate IV.

[56]) a. a. O. [16]), S. 69 und plate IV.

[57]) a. a. O. [16]), S. 70 und plate IV.

[58]) Vgl. [54]) mit Darwins ursprünglicher Figur in Math. Annalen 51 (1899), Tafel (am Schluß des Bandes), Fig. 8.

bedingt hat[59]). Zu deren Behebung hatte Darwin[60]) auf Anregung von Burrau[61]) in einer späteren Arbeit[62]) die Stoßbahn der ersten Phase berücksichtigt und für die zweite Phase (das ist für $\mu = \frac{1}{2}$ die Entwicklungsperiode VII bis X*) gewisse Ansätze gemacht. Was nun die dritte Phase (XI bis XV*) der Entwicklung anlangt, so enthält diese Darwins Satellitenfamilie B und *außerdem* Darwins Satellitenfamile C, was durch Darwin — unter ersichtlichem Einfluß eines irreführenden Poincaréschen Satzes[63]) — nur deshalb nicht bemerkt worden ist[64]), weil die Jacobische Konstante in der dritten Phase (und übrigens auch in der ersten; vgl. aber weiter unten) sich nicht monoton verhält, sondern zwischen den Schlußstoßbahnen der zweiten und der dritten Phase ein Maximum hat. Diejenige Hälfte der dritten Phase, die vor dem Maximum der Jacobischen Konstante in der dritten Phase liegt [vgl. (133)], nennt nun Darwin Familie B, die zweite Hälfte der dritten Phase aber Familie C, endlich die „Grenze" zwischen den beiden Familien, die zum relativen Maximum der Jacobischen Konstante gehört, eine solche Bahn, von welcher die beiden Familien „abzweigen". In Wirklichkeit handelt es sich freilich um eine ganz künstliche Einteilung, an der der Aufbau jeder Genealogie scheitern mußte[65]). Man muß einfach die Jacobischen Konstanten in der Familie B monoton wachsend, in C aber monoton abnehmend ordnen (bei Darwin ist die Anordnung stets die monoton abnehmende), um einzusehen, daß von einer „Abzweigung" ebensowenig die Rede ist wie von einer Bestätigung jenes von Darwin[66]) erwähnten Poincaréschen Satzes[67]). Eine ausführliche Analyse der Gruppe g soll nicht in dieser Mitteilung gegeben werden [ich bemerke in diesem Zusammenhange nur, daß (133) und (24) noch keinesfalls vermuten lassen, daß die einfachen Bahnen XIII bis XIV in der Hekubawelle liegen! Vgl. übrigens oben]. Was ich hier beweisen will, gelingt bereits auf Grund der Satellitenfamilie A.

---

[59]) Vgl. S. 336 der ersten Mitteilung [2]).

[60]) a. a. O. [16]), S. 139—141 und S. 169ff. und übrigens bereits (1899) a. a. O. [53]), S. 523, Fußnote.

[61]) Vgl. S. 322 der ersten Mitteilung [2]).

[62]) Aus dem Jahre 1910; wieder abgedruckt a. a. O. [60]).

[63]) Vgl. a. a. O. [2]).

[64]) S. auch noch a. a. O. [16]), S. 175.

[65]) Vgl. die durch Fischer-Petersen a. a. O. [21]), S. 19—21 veröffentlichten, für die ganze Theorie grundlegenden Ausführungen von Strömgren, ferner a. a. O. [18]), S. 16—18. Diese Erkenntnis fehlt auch in der letzten Veröffentlichung von Darwin [vgl. insb. a. a. O. [16]), S. 175—176], trotzdem auch Darwin gelegentlich mit „Charakteristiken" (im Sinne der Kopenhagener Arbeiten) arbeitet, vgl. a. a. O. [16]), S. 169—170. Die Darwinsche Bemerkung über die Stellung der Hillschen Spitzbahn [a. a. O. [16]), S. 71] kann auch nicht richtig sein.

[66]) a. a. O. [16]), S. 41.

[67]) Vgl. a. a. O. [2]).

Die [oben (S. 356) durch D bezeichneten] Darwin-Jacobischen Konstanten der von Darwin berechneten Bahnen der Satellitenfamilie A sind die folgenden[54]) [NB. Darwin beginnt die Rechnungen nicht in dem trivialen massennahen, d. h. „kreisähnlichen" Anfangsstück der Gruppe, also nicht bei ganz großen Werten von D, sondern grundsätzlich[68]) erst dort, wo es keine umschließende Nullgeschwindigkeitskurve mehr gibt]:

(134) $D(\mu = \frac{10}{11})$: 40,5, 40,25, 40,2, 40, 39,8, 39,5, 38,85.

Diese Serie entspricht also der ersten Zeile von (133) [wobei wir die drei Bahnen I bis III, die noch „kreisähnlich" sind und insbesondere noch brauchbare Hillsche Reihenentwickelungen haben[49]), nicht anschreiben wollen]:

(135) $K(\mu = \frac{1}{2})$: 14,95, 15,35, 13,94.

Insbesondere sind $D = 33{,}85$ und $K = 13{,}94$ die ersten Stoßbahnen. Reduziert man diese K-Werte mittels (25) auf D-Werte, so geht (135) in

(136) $D(\mu = \frac{1}{2})$: 42,04, 43,12, 39,26

über. Vergleicht man (134) mit (136) oder die entsprechenden Figuren[69]) miteinander, so ersieht man, daß die Entwicklung der Gruppe in ihren nicht mehr kreisähnlichen Bahnen bei Darwin ($\mu = \frac{10}{11}$) *langsamer* vor sich geht als bei Strömgren-Burrau ($\mu = \frac{1}{2}$); z. B. die Jacobische Konstante der Stoßbahn ist in (136) gleich 38,85, in (136) aber noch 39,26. Dies ist aber im Rahmen des Störungsprinzips, d. h. im Rahmen der Sortenmannigfaltigkeit ganz unverständlich [einmal handelt es sich um den $\frac{10}{11}$-Querschnitt, das andere Mal um den $\frac{1}{2}$-Querschnitt der Mondwelle]. Denn es wäre zu erwarten, daß die Entwicklung der Gruppe desto schneller vor sich geht, je größer der Massenprozentsatz ist, denn dann sind auch die Störungen größer.

Dieses Paradoxon behebt sich nun wieder dann und nur dann, wenn man den Rahmen der Störungsproblematik verläßt [d. h. wenn man beachtet, daß die Bahnen (134), (136) gar nicht auf den Querschnitten der Mondwelle, sondern bereits oberhalb der Massenschranke liegen müssen]. Dann und nur dann verliert nämlich die obige Schlußweise, daß nämlich zu $u = \frac{10}{11}$ eine größere Störung, also eine rapidere Gruppenentwicklung als zu $u = \frac{1}{2} < \frac{10}{11}$ gehören muß, ihren Inhalt: Die „Störungen" sind gar nicht definiert, man befindet sich vielmehr bereits oberhalb der Massenschranke. In einem solchen Gebiete verliert also die Entwicklung der Gruppe jeden auch qualitativen Kontakt mit den Begriffsbildungen der Problematik der Sorten.

---

[68]) a. a. O. [16]), S. 53.

[69]) a. a. O. [52]), Fig. 5 bzw. a. a. O. [16]), plate IV, fig. 1.

Die zu (133) gehörigen C-Werte sind nach (24) die folgenden:

$$C(\mu=\tfrac{1}{2}):\ 50,\ 16,\ 6,\ 3{,}14,\ 3{,}83,\ 3{,}49. \tag{137}$$

Die Jacobische Konstante hat also noch vor der ersten Stoßbahn $C = 3{,}49$ der Gruppe g ein Minimum. Dies ist nach den Hillschen Rechnungen[70]) bereits in dem parallaxenlosen Problem der Fall. Bei Darwin ist diese Erscheinung dennoch nicht ersichtlich [vgl. (134)], sei es, weil sie für $\mu = \frac{10}{11}$ nicht existiert (was mit Rücksicht auf die Figuren nicht eben wahrscheinlich ist), sei es, weil die von Darwin berechnete Serie (134) unvollständig ist. Aus (137) folgt jedenfalls [vgl. die Tabelle auf S. 369], daß das relative Minimum von $C(\mu=\frac{1}{2})$ zwischen $C = +\infty$ und der Stoßbahn $C = 3{,}49$ in der Gegend des Lückenzentrums $(C)_{m=1}$ des Hekubatypus liegt, was freilich unter keinen Umständen zu diesbezüglichen Schlüssen berechtigt. Es folgt hingegen bereits aus dem nicht monotonen Gange der Jacobischen Konstante, daß die Kopenhagener Gruppe g recht bald oberhalb der Massenschranke zu liegen kommt, da doch C auf jedem Querschnitt der Mondwelle eine monotone Funktion des Ortes ist[71]).

§ 9.

Die übrigen Gruppen mit Sortenursprung.

In diesem Paragraphen sollen einige Bemerkungen über die Gruppen f, l, m und k zusammengestellt werden[72]); die vollständige Theorie kann ich erst später entwickeln, z. B. für l und k erst nach der Librationsgenealogie.

So wie die erste Relation (118) das astronomisch selbstverständliche Anfangsstück der Gruppe g bei jedem μ sichergestellt hat (§ 7), leisten die übrigen drei Relationen (118) offenbar dasselbe für die drei Gruppen f, l, m (und zwar wieder bei jedem μ), wie es bereits in den beiden letzten Zeilen der Tabelle (S. 369) angegeben worden ist. Mit Rücksicht auf die erste Mitteilung[2]) können daher die Kopenhagener numerischen Ergebnisse $(\mu=\frac{1}{2})$ auch bei diesen drei Gruppen als mathematisch fundiert gelten. — Wie aus den Fourierreihen hervorgeht [vgl. S. 391, S. 381], sind alle vier Gruppen g, f, l, m in ihrem massennahen bzw. massenfernen Anfangsstück symmetrisch in bezug auf die Syzygienachse, und zwar bei

---

[70]) a. a. O. [14]).

[71]) An sich besteht freilich die Möglichkeit, daß man bei sortenmäßiger Fortsetzung, aber nicht auch bei gruppenmäßiger Fortsetzung zu einer effektiven Verzweigungsstelle gelangt oder umgekehrt, vgl. a. a. O. [50]), aber diese Schwierigkeit besagt wieder nur die Unnatürlichkeit jeder sortenmäßigen Konstruktion.

[72]) Vgl. a. a. O. [12]), S. 35—36, S. 41—46 oder a. a. O. [18]), S. 15, S. 19—24, oder a. a. O. [11]), S. 985—987.

jedem μ (l und m im Falle $\mu = \frac{1}{2}$, aber auch nur in diesem Falle, haben außerdem auch die η-Achse zur Symmetrieachse), so daß ξ eine gerade, η eine ungerade Funktion von t ist, d. h. die Sinus- bzw. Kosinuskoeffizienten verschwinden identisch in der Jacobischen Konstante C. Da sie aber analytische Funktionen von C sind, so gilt dies nicht nur in dem Anfangsstück, sondern überhaupt. D. h., alle Bahnen der vier Gruppen g, f, l, m sind symmetrisch in bezug auf die Syzygienachse und müssen daher diese mindestens zweimal orthogonal passieren, wobei ein orthogonales Passieren ausnahmsweise eventuell durch eine Spitze ersetzt ist, nämlich bei den Ejektionsbahnen (sowie a priori auch bei denjenigen Bahnen der Gruppe, die die zugehörige Nullgeschwindigkeitskurve auf der Syzygienachse treffen, doch sind keine solchen Bahnen bekannt geworden). — Ein Vergleich des Kopenhagener Materials ($\mu = \frac{1}{2}$) mit dem Darwinschen ($\mu = \frac{1}{11}$ bzw. $\mu = \frac{10}{11}$), so wie bei der Gruppe g (§ 8), ist bei den drei synodisch retrograden Gruppen f, l, m leider nicht möglich, da Darwin sich zunächst[73]) grundsätzlich[74]) auf synodisch direkte Bahnen beschränkt hatte. Im übrigen liegen die Verhältnisse bei f, l, m analog wie bei g, so daß wir uns ganz kurz fassen wollen.

Die Gruppe f[75]) beginnt, ebenso wie g, mit massennahen kreisähnlichen Bahnen, die aber in g (siderisch und synodisch) direkt, in f (siderisch und synodisch) retrograd sind. f wird durch dieselben Fourierreihen (131) wie g festgelegt, nur daß $\sqrt{C}$ ($C \to +\infty$) in (132) bei g positiv, bei f negativ zu wählen ist[76]); jedenfalls ist f (bei $C = +\infty$) die analytische Fortsetzung von g[77]) (vgl. auch Fig. 2 bei $m = 0$). Es versteht sich also von selbst, daß der direkte Existenzbeweis[78]) des Anfangsstückes von f von dem entsprechenden Beweis für g sich nicht unterscheidet. Hier sind auch die anderweitig gewonnenen Ergebnisse von Birkhoff[79]) über retrograde Bahnen zu erwähnen, die freilich auch in dem obigen Rahmen erhalten und — mit Rücksicht auf das oben über orthogonales Achsenpassieren Gesagte — offenbar auch verschärft werden könnten. — Die Gruppe f enthält keine Lückenzentra (vgl. die Tabelle auf S. 369) und die Entwicklung der Gruppe f erfolgt viel langsamer als diejenige von g, d. h. die

---

[73]) Vgl. allerdings a. a. O. [16]), S. 155—158.

[74]) a. a. O. [16]), S. 53.

[75]) Über diese Gruppe s. C. Burrau und E. Strömgren, Astr. Nachr. **202** (1916), S. 185—199 = Publ. Købn. Observ. Nr. 23 und J. P. Möller, Astr. Nachr. **233** (1928), S. 81 ff. = Publ. Købn. Observ. Nr. 60.

[76]) Vgl. a. a. O. [49]) mit J. Fischer-Petersen, Astr. Nachr. **202** (1916), S. 201—204 = Publ. Købn. Observ. Nr. 23.

[77]) Vgl. a. a. O. [2]).

[78]) Vgl. oben S. 391.

[79]) a. a. O. [8]), § 18.

Bahnen von f sind, im Gegensatz zu g, bei noch relativ ziemlich kleinen Werten von C „kreisähnlich“[80]. Der Grund dieser langsameren Entwicklung wird aus den Figuren auf S. 381ff. unmittelbar ersichtlich, wenn man dabei korrespondierende Abszissengebiete unter Beachtung der Tabelle (S. 369) miteinander vergleicht. Daß das unterhalb der Massenschranke gelegene Sortengebiet auch von f bald verlassen wird, folgt wieder bereits aus dem Gange der Jacobischen Konstanten[75]), den ich also gar nicht wiedergeben will. Der natürliche Abschluß besteht nach Möller[75]) darin, daß die Bahndimension unendlich wird, nachdem die Gruppe eine Folge von störungstheoretisch erreichbaren Stoßbahnen passiert hat.

Auch die Gruppen l[81]) und m[82]) (von zentrifugalem kreisähnlichem Anfang: vgl. S. 380 oben), deren dynamisch selbstverständliches Anfangsstück aus massenfernen großen Bahnen besteht und wieder auf Grund der Methode der unbestimmten Koeffizienten angesetzt[83]) und diese mit meiner Methode[84]) legalisiert werden kann (vgl. S. 391 oben), sind (diesmal bei $C = \pm\infty$) analytische Fortsetzungen voneinander[85]) und ebenfalls nicht geschlossen. Der natürliche Abschluß von l[81]) erfolgt durch Unendlichwerden der Periode und nur der Periode, der natürliche Abschluß von m[82]) durch Unendlichwerden der Jacobischen Konstante und nur der Jacobischen Konstante. Sowohl l als auch m ragen aus der Sortenmannigfaltigkeit wieder hinaus, wie das wieder bereits aus dem Gange der Jacobischen Konstanten, bei l sogar aus der Existenz von Lückenzentren (vgl. die Tabelle auf S. 369) hervorgeht; die Kopenhagener Werte der Jacobischen Konstanten[81]) [82]) sollen daher gar nicht detailliert werden. — Weder l noch m enthält eine Stoßbahn (auch m nicht, da eine feste Abschlußbahn gar nicht existiert). Da beide Gruppen mit *großen* gestörten Kreisen beginnen, so sind sie anfangs synodisch gewiß retrograd (vgl. auch die Tabelle auf S. 369), und sie bleiben es bis zum natürlichen Abschluß, da es doch keine Stoßbahn gibt, durch die die Indikatrix umgekehrt werden könnte. l ist siderisch direkt, m siderisch retrograd (beim Gruppenursprung). — Da beide Gruppen synodisch retrograd sind, so ist die Existenz synodisch direkter, beide Massen

---

[80]) Vgl. die Figur bei Burrau-Strömgren, a. a. O. [75]), mit der Fig. 5 bei Burrau-Strömgren, a. a. O. [52]).

[81]) Über diese Gruppe s. E. Strömgren und J. Fischer-Petersen, Astr. Nachr. 207 (1918), S. 1—10 = Publ. Købn. Observ. Nr. 30, ferner E. Strömgren, Seeliger-Festschrift, 1924 = Publ. Købn, Observ. Nr. 47, S. 6ff. und S. 13, endlich E. Strömgren, Mitteilungen der Kopenhagener Akademie X. 9 (1930) = Publ. Købn. Observ. Nr. 69.

[82]) Über diese Gruppe s. J. P. Möller, Astr. Nachr. 221 (1924), S. 81—90 = Publ. Købn. Observ. Nr. 48.

[83]) P. Pedersen, Astr. Nachr. 207 (1918), S. 9—16 = Publ. Købn. Observ. Nr. 30.

[84]) a. a. O. [48]); vgl. nämlich [35]).

[85]) Vgl. a. a. O. [2]), S. 338.

umschließender Massen im Rahmen der Störungsproblematik der Sorten durchaus unverständlich. Dem entspricht die Moultonsche Behauptung, wonach es keine synodisch direkte, beide Massen umschließende periodische Bahnen geben kann[86]) — was freilich durch die Kopenhagener Gruppe k widerlegt wird: Herr Moulton übersieht die Grenzen der Störungsproblematik.

Bei der Gruppe k[87]), die ganz und gar außerhalb der Sortenmannigfaltigkeit liegt und daher in dieser Mitteilung nur mit Rücksicht auf den Moultonschen Irrtum gestreift werden soll, scheint die Kopenhagener Wahl des Massenprozentsatzes ($\mu = \frac{1}{2}$) eine wesentlichere Rolle zu spielen als bei den übrigen Gruppen. Die Gruppe k ist übrigens ebenfalls nicht geschlossen, sie erleidet vielmehr in den *beiden* Richtungen, d. h. sowohl beim „Anfangsstück" als auch beim „Endstück" (bei Strömgren mit k_1 bzw. k_2 bezeichnet) einen natürlichen Abschluß, der beidemal durch das wegen der (für $\mu = \frac{1}{2} > \frac{1}{26}$ instabilen) äquilateralen Librationspunkte unbegrenzte Anwachsen der Periode, also asymptotisch bedingt ist, und Strömgrens anschauliche Bemerkung[88]), wonach die Bahnen der Gruppe k ein in sich geschlossenes Ganze bilden, ist nur in diesem Sinne, nicht aber etwa so zu verstehen, daß die Gruppe k (so etwa wie die Librationsgruppe b) im Sinne der ersten Mitteilung[89]), d. h. in dem eigentlichen Strömgrenschen Sinne in sich geschlossen ist; k_1 und k_2 enthalten vielmehr nicht dieselben Bahnen, es handelt sich also nicht um eine in sich schließende analytische Mannigfaltigkeit (des näheren vgl. S. 347 der ersten Mitteilung). — Die (k_1 mit k_2 verbindende) einzige Stoßbahn von k ist von Herrn Moulton[90]) im Jahre 1920 veröffentlicht worden (die Moultonschen *C*-Werte sind durch 4 zu multiplizieren, um die Kopenhagener *K*-Werte zu erhalten), nachdem diese Stoßbahn viel genauer und samt der Fourierreihe sowie weiteren Bahnen der Gruppe k durch Strömgren und Burrau[87]) im Jahre 1916 veröffentlicht worden ist und nachdem Lous[87]) (1919) im Anschluß daran auch weitere Bahnen der Gruppe k berechnet hat. Aus den Kopenhagener Rechnungen[87]) geht unmittelbar hervor, daß die sich auf die Fortsetzung der Stoßbahn beziehenden Vermutungen von Herrn Moulton[91]) ganz falsch sind, trotzdem sie erst im Jahre 1920 erschienen.

[Auf weitere Untersuchungen von Herrn Moulton, insbesondere seine

[86]) F. R. Moulton, Periodic Orbits, Washington 1920.

[87]) Über die Gruppe s. C. Burrau und E. Strömgren, Astr. Nachr. **203** (1916), S. 277—284 = Publ. Købn. Observ. Nr. 26, ferner K. Lous, Astr. Nachr. **209** (1919), S. 241—254 = Publ. Købn. Observ. Nr. 32, endlich E. Strömgren, Seeliger-Festschrift. 1924 = Publ. Købn. Observ. Nr. 47, S. 10—11.

[88]) a. a. O.[11]), S. 238.

[89]) a. a. O.[2]), S. 347.

[90]) a. a. O.[86]), S. 516.

[91]) a. a. O.[86]), S. 517.

übrigen genealogischen Behauptungen in den Periodic Orbits[92]), soll in den späteren Mitteilungen auch des näheren eingegangen werden, z. B. im Zusammenhang mit den Librationen und den Doppelbahnen bzw. bei den sogenannten Verzweigungen der Darwinschen Familien. Hingegen sind die Sätze von Herrn Moulton (1920) über Spitzen und Schleifen[93]) ganz richtig, sind aber von Strömgren im Jahre 1906 veröffentlicht worden[94]), während die den Umlaufssinn umkehrende Rolle der Stoßbahnen von Burrau[95]) erkannt worden ist.]

* * *

Es versteht sich von selbst (vgl. S. 385, S. 386), daß z. B. auch die μ-Querschnitte $\mathfrak{W}_j(\mu)$ und $\mathfrak{Z}_n(\mu)$ Gruppenstücke ergeben, für welche der Abschlußsatz der ersten Mitteilung gilt. Worin dieser Abschluß besteht, ist aber nicht bekannt, da keine Einzeluntersuchungen vorliegen. Jedenfalls werden diese Gruppen, und zwar bereits in ihrem noch in der Sortenmannigfaltigkeit liegenden Anfangsstück, z. B. bei wachsendem $|p-q$, wobei $n=p/q$, $(p,q)=1$ gesetzt ist, der Bahngestalt nach immer komplizierter, so daß sie bereits aus diesem Grunde außerhalb des Strömgrenschen Programms lagen, und andererseits ist sowohl bei $\mathfrak{W}_j(\mu)$ als auch bei $\mathfrak{Z}_n(\mu)$ zu beachten, daß diese Gruppenstücke vielleicht nur bei sehr kleinem μ existieren, da hier Analogien zu (118) weder bekannt noch plausibel sind.

§ 10.

Die isoperiodischen Sorten.

Die sortentheoretischen Untersuchungen von Poincaré über das restringierte Dreikörperproblem sind in verschiedenen Arbeiten zerstreut; es kommen hier I: der erste Band der Méth. Nouv., II: der dritte Band der Méth. Nouv., III: die Genealogie-Note in den Bull. Astr., endlich IV: die posthume topologische Arbeit in Betracht, während Poincarés ältere Untersuchungen in den Méth. Nouv. aufgegangen sind. IV kann hier beiseite gelassen werden, da darin das Existenzproblem im kleinen bei den μ-Strecken zweiter Sorte, d. h. die Existenz wenigstens der einzelnen μ-Strecken zweiter Sorte mit Rücksicht auf I, II, III als erledigt betrachtet wird, was aber, wie wir weiter unten sehen werden, nicht der Fall war. II käme nur dann in Betracht, wenn Poincarés Ausführungen über den Zusammenhang der zweiten Sorte mit dem zweiten Genre sowie die Theorie des zweiten Genres in II eine exakte Begründung gefunden hätten. Dies ist freilich, wie bereits in der ersten Mitteilung[2]) erwähnt wurde, z. Z. nicht der Fall. In III werden

[92]) a. a. O. [86]), S. 485—524.
[93]) a. a. O. [86]), S. 495 ff.
[94]) E. Strömgren, Astr. Nachr. **174** (1906), S. 33 ff.
[95]) Vgl. S. 322 der ersten Mitteilung [2]).

keine Existenzbeweise gegeben, es wird vielmehr auf Grund von I angenommen, daß die isoenergetischen μ-Strecken (in III handelt es sich ausdrücklich um nicht isoperiodische Sorten) mit Ausnahme des Falles $e = 0$, $m = j$ existieren. Demnach kommt für uns nur I in Betracht. In I handelt es sich nicht um die isoenergetischen, sondern ausdrücklich [96]) um die isoperiodischen Sorten des restringierten Dreikörperproblems und für den isoenergetischen Fall wird nur der allgemeine Poincarésche Störungsansatz angegeben [97]). — Die in der Genealogie-Note [5]) (Bull. Astr.) als existierend vorausgesetzten einzelnen isoenergetischen μ-Strecken sind für die erste Sorte zuerst von Herrn Levi-Civita sichergestellt worden [98]).

Wir gehen nun zu I und damit zu den isoperiodischen Sorten über. Über die isoperiodische erste Sorte wird in I bewiesen, daß der Punkt $(m, e = 0)$ der a priori zulässigen Mannigfaltigkeit $\mu = 0$ jedenfalls dann (isoperiodisch) fortsetzbar ist, wenn $m \neq j$ angenommen wird. Die weitere Literatur der isoperiodischen ersten Sorte ist in der vorläufigen Mitteilung [1]) besprochen worden. Es zeigt sich, daß alle oben (§ 5) über die isoenergetische erste Sorte zusammengestellten Sätze auch für die isoperiodische erste Sorte richtig sind, das heißt, die oben für die isoenergetische Massenschranke M angegebenen Beziehungen bleiben im Falle $e = 0$ richtig, wenn an Stelle von M das auf S. 377 definierte Funktionszeichen N geschrieben wird. Was die isoperiodischen μ-Strecken zweiter Sorte anlangt, so nimmt Poincaré an, daß ihre Existenz in I bewiesen worden ist. Stäckel [99]) hat indessen in einer nachgelassenen Note darauf hingewiesen, daß die Poincarésche Schlußweise einen Fehler enthält. Ob die isoperiodische zweite Sorte dennoch existiert oder aber das Poincarésche Schlußresultat (d. h. der Existenzsatz) fallen gelassen werden muß, bleibt bei Stäckel dahingestellt. In der vorläufigen Mitteilung [100]) (in welcher ich die Stäckelsche Note nicht zitiert habe, da sie mir leider entgangen war) habe ich nun gezeigt, daß der isoperiodische Poincarésche Existenzsatz, eben für hinreichend kleine Werte der Eigenexzentrizität, gewiß fallen gelassen werden muß, was mit Rücksicht auf anderweitige Rechnungen von Hill und Wilkens eigentlich vorauszusehen war.

* * *

An dieser Stelle ist auch eine Jugendarbeit von Schwarzschild [7]) zu erwähnen, welche ebenfalls die zweite Sorte behandelt. Die μ-Strecken sind dabei auf eine nicht isoperiodische Weise angesetzt, derart, daß die

[96]) a. a. O. [3]), S. 97 ff. bzw. S. 139 ff.
[97]) a. a. O. [3]), S. 89 ff.; vgl. auch a. a. O. [4]), S. 207 ff.
[98]) Vgl. S. 379 oben.
[99]) P. Stäckel, Jahresber. d. Deutschen Math.-Ver. 28 (1919), S. 180–181.
[100]) a. a. O. [1]), § 9.

Schwarzschildsche „Normierung“ (vgl. S. 374 oben) mit der isoenergetischen durchaus äquivalent ist. Es scheint also, daß die Existenz von isoenergetischen μ-Strecken bereits vor den Birkhoffschen Arbeiten feststand. Dies ist aber nicht der Fall, bei Schwarzschild wird nämlich die Diskussion der „Verzweigungsgleichung“ (d. i. diesmal die Perihelbedingung), die den springenden Punkt darstellt, auf eine nicht korrekte Weise behandelt. Dies habe ich bereits in der vorläufigen Mitteilung erwähnt. Auf S. 23 [loc. cit. [7])] wird nämlich *nicht das gestörte* Perihel eingeführt, *und zwar auch dann nicht*, wenn (im letzten Absatz von S. 23) von $\mu = 0$ zu $\mu \neq 0$ übergegangen wird, was nur durch den Umstand verschleiert ist, daß Schwarzschild den zu der Epoche gehörigen Wert des Perihels nicht in seiner Abhängigkeit von μ anschreibt und damit den Buchstaben ω in zwei wesentlich verschiedenen Bedeutungen verwendet. Dies ist wohl auch der Grund, daß der Sachverhalt bisher unbemerkt bleiben konnte. — Daß die Schwarzschildsche Schlußweise nicht stichhaltig sein kann, folgt übrigens schon daraus, daß Schwarzschild die μ-Strecke bei jeder Kommensurabilität und bei jeder Eigenexzentrizität $\geqq 0$, also auch für $e = 0$, $m = j$ konstruieren kann, was doch nach (114) nicht zutreffen kann. [Es ist weniger wichtig und soll nur der Vollständigkeit halber erwähnt werden, daß in dem Fall einer nicht kleinen Eigenexzentrizität die Betrachtungen schon deshalb nicht gelten, weil die benutzte Entwicklung der Störungsfunktion ungültig ist.] — Die Vereinfachung, die Herr Charlier a. a. O. an der Schwarzschildschen Diskussion der Perihelbedingung angebracht hat, sowie die Charliersche Theorie der isoenergetischen *und* isoperiodischen zweiten Sorte werde ich demnächst an anderer Stelle besprechen.

Jedenfalls sieht man jetzt, weshalb in dem vorliegenden Aufsatz das Eigenperihel der Epoche von vornherein gleich Null gesetzt worden ist (S. 367). Bezeichnet nämlich e' die Exzentrizität von Jupiter, so ist in dem restringierten Problem, mit dem wir uns ausschließlich beschäftigt haben, $e' = 0$, und auch bei Schwarzschild ist $e' = 0$. Mit Rücksicht auf den Stand der Dinge bei Schwarzschild hätte es keinen Sinn, den sowieso recht verzweigten Sortenraum durch Einführung eines nicht verschwindenden epochalen Eigenperihels (also einer vierten „Dimension“) noch wesentlich zu komplizieren. Im Falle des reduzierten Dreikörperproblems ($e' \neq 0$), worauf sich Rechnungen von Hill und Wilkens (und gewisse neuere Ansätze von Heinrich) beziehen [6]), wäre freilich diese Vereinfachung keinesfalls erlaubt gewesen.

The Johns Hopkins University, den 10. Januar 1931.

(Eingegangen am 4. März 1931.)

Some properties of fractional integrals. II.

By

G. H. Hardy in Oxford and J. E. Littlewood in Cambridge.

1. Introduction.

1.1. This paper does not demand much preliminary explanation, since it is essentially an extension to the complex field of our earlier paper (7)[1]) with the same title.

In 7 we were concerned with functions of a real variable belonging to one or other of the "Lebesgue classes" L^p, where $p \geqq 1$. Here our functions are analytic (or harmonic) functions of $z = re^{i\theta}$, where $0 \leqq r < 1$; and a function is said to belong to L^p when the mean M_p defined below is bounded. The scope of our analysis is then enlarged in two ways. In the first place we have to consider the range $p > 0$ instead of the range $p \geqq 1$. Secondly we have to consider not only when a mean M_p is bounded but also when it is of a given order $(1-r)^{-\alpha}$. The chief novelties of the paper arise from the first extension, the case $p < 1$ being decidedly more difficult than the case $p \geqq 1$.

The contrast between these two cases becomes much more striking when we consider the analogous problems connected with a pair of conjugate harmonic functions u, v, since then it affects the results as well as the methods of proof. These problems we reserve for a separate paper.

The means M_p and the classes L^p.

1.2. Suppose that

$$f(z) = f(re^{i\theta}) \qquad (r \geqq 0)$$

[1]) See the list of references at the end. We number our theorems in succession to those of 7.

is an analytic function of z regular for $r = |z| < 1$; and that $p > 0$. Then we write

$$M_p = M_p(f) = M_p(r) = M_p(r, f) = \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(re^{i\theta})|^p d\theta\right)^{\frac{1}{p}}.$$

The following properties of these means are well known.

(1) M_p is an increasing function of p.

(2) $\log M_p^p$ is a convex function of p.

(3) M_p tends to

$$M = M(f) = M(r) = M(r, f),$$

the maximum modulus of $f(z)$, when $p \to \infty$.

(4) M_p tends to

$$M_0 = \exp\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \log|f(re^{i\theta})|\, d\theta\right)$$

when $p \to 0$.

(5) M_p is an increasing function of r.

(6) $\log M_p$ is a convex function of $\log r$[2]).

If, for any p, $M_p(f) < K$, where K is independent of r, we say that $f(z)$ *belongs to the* (complex) *class* L^p or, more shortly, that $f(z)$ *is* L^p.

We may frame similar definitions for harmonic functions. Suppose that

$$u(z) = u(re^{i\theta}) = u(r, \theta)$$

is a complex harmonic function regular for $r < 1$. Then we write

$$M_p = M_p(u) = M_p(r) = M_p(r, u) = \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |u(re^{i\theta})|^p d\theta\right)^{\frac{1}{p}},$$

and define the classes L^p as before. These means possess properties analogous to (1)—(6) when $p \geqq 1$. We shall not be concerned in this paper with means of harmonic functions in which $p \leqq 1$.

1.3. We collect here, in the form of a comprehensive lemma, a number of known theorems concerning power series, Fourier series, and boundary functions.

We write

$$P(r, \theta) = \frac{1 - r^2}{1 - 2r\cos\theta + r^2}, \qquad Q(r, \theta) = \frac{2r\sin\theta}{1 - 2r\cos\theta + r^2}.$$

---

[2]) For (1)–(4) see Littlewood (**11**), and Hausdorff's paper there referred to; for (5) and (6) see Hardy (**2**), Landau (**10**), Pólya und Szegö (**12**, p. 143, 329), F. Riesz (**13**, **14**).

Lemma A. (1) *If* $p > 0$, $f(z) = f(re^{i\theta})$ *is regular for* $r < 1$, *and* $M_p(f) \leqq C$, *then* $f(z)$ *tends radially to a limit*

$$f(e^{i\theta}) = F(\theta)$$

for almost all θ, *and*

$$M_p(F) = \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |F(\theta)|^p \, d\theta\right)^{\frac{1}{p}} \leqq C.$$

(2) *If further* $p \geqq 1$, *then* $f(z)$ *is the Poisson integral, and the Cauchy integral, of* $F(\theta)$; *i. e.*

$$f(z) = \frac{1}{2\pi}\int_{-\pi}^{\pi} P(r, \varphi - \theta)\, F(\varphi)\, d\varphi = \frac{1}{2\pi i}\int_{-\pi}^{\pi} \frac{F(\varphi)\, de^{i\varphi}}{e^{i\varphi} - z}.$$

(3) *If* $p > 0$, $F(\theta)$ *is a function of* L^p *whose Fourier series is a Fourier power series*

$$\sum_0^\infty c_n e^{ni\theta},$$

and $f(z) = \sum c_n z^n$, *then* $M_p(f) \leqq M_p(F)$; *and, when* $p \geqq 1$, $f(z)$ *and* $F(\theta)$ *are related as in* (1) *and* (2) *above.*

(4) *If* $p > 1$, $u(z) = u(r, \theta)$ *is a (real or complex) harmonic function regular for* $r < 1$, *and* $M_p(u) \leqq C$, *then* $u(r, \theta)$ *tends radially to a limit*

$$u(1, \theta) = U(\theta)$$

for almost all θ, *and* $M_p(U) \leqq C$. *Also* $u(r, \theta)$ *is the Poisson integral of* $U(\theta)$.

(5) *If* $p > 1$, $U(\theta)$ *is a function of* L^p *whose Fourier series is*

$$\sum_{-\infty}^{\infty} c_n e^{ni\theta} \quad \left[\text{or } \frac{1}{2}a_0 + \sum_1^\infty (a_n \cos n\theta + b_n \sin n\theta)\right],$$

and

$$u(r, \theta) = \sum_{-\infty}^{\infty} c_n r^{|n|} e^{ni\theta} \left[\text{or } \frac{1}{2}a_0 + \sum_1^\infty (a_n \cos n\theta + b_n \sin n\theta) r^n\right],$$

then $M_p(u) \leqq M_p(U)$, *and* $u(r, \theta)$ *and* $U(\theta)$ *are related as in* (4) *above.*

(6) *If* $p > 1$ *and* u *is a real harmonic function for which* $M_p(u) \leqq C$, *then there are positive harmonic functions* u_1 *and* u_2 *such that*

$$u = u_1 - u_2, \qquad M_p(u_1) \leqq C, \qquad M_p(u_2) \leqq C.$$

For proofs and explanations of these propositions we may refer to F. and M. Riesz (**15**), F. Riesz (**13**), M. Riesz (**16**), Hardy and Littlewood (**5**). It is to be observed that here, and throughout the paper, we use C (as

opposed to K) for a number which preserves its identity throughout a theorem and its proof.

The various clauses of Lemma A indicate important differences between the ranges $p > 1$ and $p \leqq 1$. The most striking of these appears in the well known theorem of M. Riesz. If $p > 1$, u is L^p, and v is a conjugate of u, then v and $f(z) = u + iv$ are L^p [3]). It is familiar that this is false for $p = 1$ [4]), and there are simple examples which show that it is also false for $p < 1$. When $p < 1$, however, the proposition is false in a much more comprehensive sense, to which we shall return in the paper mentioned at the end of § 1.1.

We shall also make much use of

Lemma B [5]). *If $p > 0$, $f(z)$ is regular for $r < 1$, and $M_p(f) \leqq C$, then*

$$f = f_1 + f_2,$$

where (a) f_1 and f_2 are regular for $r < 1$, (b) f_1 and f_2 have no zeros for which $r < 1$, and (c)

$$M_p(f_1) \leqq 2C, \quad M_p(f_2) \leqq 2C.$$

The result is true in the limiting case $p = \infty$, M_p being replaced by M.

2. Some theorems concerning the means M_p.

2.1. Theorem 27. *If $p > 0$, $a \geqq 0$,*

$$M_p(f) \leqq C(1-r)^{-a}, \tag{2.11}$$

and $q > p$, then

$$M_q(f) \leqq KC(1-r)^{-a-\frac{1}{p}+\frac{1}{q}}, \tag{2.12}$$

$$M(f) \leqq KC(1-r)^{-a-\frac{1}{p}}, \tag{2.13}$$

where $K = K(p, a)$.

If $a = 0$, then also

$$M_q(f) = o\left((1-r)^{-\frac{1}{p}+\frac{1}{q}}\right), \tag{2.14}$$

$$M(f) = o\left((1-r)^{-\frac{1}{p}}\right), \tag{2.15}$$

when $r \to 1$.

---

[3]) M. Riesz, 16.

[4]) The example being

$$u = P(r, \theta), \quad v = Q(r, \theta).$$

[5]) See 4, p. 207.

When $p \geqq 1$, all this is proved in our paper 8 [6]), so that we need only consider the case $p < 1$.

(i) Suppose first that $f(z)$ has no zeros in the unit circle. Then we may write $f = \varphi^k$, where $k = k(p)$, $kp \geqq 1$; so that

$$M_{kp}(\varphi) \leqq C^{\frac{1}{k}}(1-r)^{-\frac{a}{k}},$$

and therefore

$$M_{kq}(\varphi) \leqq KC^{\frac{1}{k}}(1-r)^{-\frac{a}{k}-\frac{1}{kp}+\frac{1}{kq}},$$

where $K = K(kp, a, k) = K(p, a)$. This is equivalent to (2.12); and (2.13), (2.14), and (2.15) follow similarly.

(ii) Passing to the general case, suppose that

$$0 < r < 1, \quad \varrho = \tfrac{1}{2}(1+r).$$

Then

$$M_p(f) \leqq C(1-\varrho)^{-a} = C_\varrho$$

for $r < \varrho$. By Lemma B, we can find f_1 and f_2 so that $f = f_1 + f_2$, $f_1 \not\equiv 0$, $f_2 \not\equiv 0$, and

$$M_p(f_1) \leqq 2C_\varrho, \quad M_p(f_2) \leqq 2C_\varrho \tag{2.16}$$

for $r < \varrho$ [7]). Applying the result of (i) to $f_1(z) = F_1(\varrho\zeta)$ in the circle $|\zeta| < 1$, we obtain

$$M_q(f_1) \leqq 2KC_\varrho\left(1-\frac{r}{\varrho}\right)^{-\frac{1}{p}+\frac{1}{q}}$$

$$= 2KC(1-\varrho)^{-a}\left(1-\frac{r}{\varrho}\right)^{-\frac{1}{p}+\frac{1}{q}} \leqq KC(1-r)^{-a-\frac{1}{p}+\frac{1}{q}}.$$

There is a similar inequality for f_2; and, since

$$M_q(f) \leqq KM_q(f_1) + KM_q(f_2),$$

we deduce

$$M_q(f) \leqq KC(1-r)^{-a-\frac{1}{p}+\frac{1}{q}},$$

which is (2.12). The proofs of the other results follow the same line.

It should however be observed that there is a difference between the logic of the arguments for the cases $a > 0$ and $a = 0$. Our argument

---

[6]) Theorem 2, pp. 623–625. The result is stated in "O" form, but an examination of the argument shows at once that K depends on p and a only.

[7]) f_1 and f_2 depend upon ϱ and are regular only for $r < \varrho$. Here and elsewhere we use the obvious extension of Lemma B to circles of any radius.

under (ii) shows in effect that "it is sufficient to prove the theorem for a function which has no zeros in the unit circle"; and this although we cannot assert that a function satisfying the conditions of the theorem is necessarily the sum of two functions which also satisfy them and have no zeros in the circle. When $a = 0$ the matter is simpler, since we can apply Lemma B directly to the unit circle, this being indeed essential, since the proof when $p \geqq 1$ involves the boundary function. Arguments of the same type will recur frequently in the sequel, and we shall not always write them out in full.

2.2. Theorem 28. *If*

$$f(z) = \sum a_n z^n$$

satisfies (2.11), *then*

$$a_n = O\left(n^{a+\frac{1}{p}-1}\right) \qquad (p \leqq 1), \tag{2.21}$$

$$a_n = O(n^a) \qquad (p > 1). \tag{2.22}$$

If $a = 0$, *we may replace* O *by* o.

If Γ is the circle $r = e^{-\frac{1}{n}}$, we have

$$|a_n| = \left|\frac{1}{2\pi i}\int_\Gamma \frac{f(z)}{z^{n+1}}\,dz\right| \leqq e M_1\left(e^{-\frac{1}{n}}\right).$$

If $p \leqq 1$ the results follow from (2.12) and (2.14). If $p > 1$, we have only to observe that $M_1 \leqq M_p$.

The examples

$$f(z) = (1-z)^{-a-\frac{1}{p}} \qquad (p \leqq 1),$$

$$f(z) = \sum 2^{an} z^{2^n} \qquad (p > 1),$$

show that the indices in (2.21) and (2.22) are the best possible.

Returning to Theorem 27, we observe that when $p \leqq 2$ we can assert more than is asserted by (2.13) and (2.15).

Theorem 29. *If* $0 < p \leqq 2$, $a \geqq 0$, *and* $f(z)$ *satisfies* (2.11), *then*

$$\mathfrak{F}(r) = \sum |a_n| r^n \leqq K C (1-r)^{-a-\frac{1}{p}}. \tag{2.23}$$

If $a = 0$, *then*

$$\mathfrak{F}(r) = o\left((1-r)^{-\frac{1}{p}}\right). \tag{2.24}$$

This is proved for $1 < p \leqq 2$ in our paper 8 [8]) (for $p > 2$ it is false). If $p \leqq 1$, it is an immediate corollary of Theorem 28.

[8]) p. 625.

3. The fractional integrals of an analytic function.

3.1. Suppose that

$$f(z) = \sum a_n z^{n+c} = z^c \sum a_n z^n = z^c g(z), \tag{3.11}$$

where c is real [9]), z^c has its principal value

$$z^c = e^{c \log z} \quad (-\pi < \Im \log z \leqq \pi),$$

and $g(z)$ is regular for $r < 1$; and that α is any real number. Then we define $f_\alpha(z)$, the integral of $f(z)$ of order α, by the formula

$$f_\alpha(z) = \sum \frac{\Gamma(n+1+c)}{\Gamma(n+1+c+\alpha)} a_n z^{n+c+\alpha}, \tag{3.12}$$

where $z^{c+\alpha}$ has its principal value. The definition obeys the formal laws

$$(f_\alpha)_\beta = (f_\beta)_\alpha = f_{\alpha+\beta},$$

and $f_\alpha(z)$ reduces to a derivative of $f(z)$ when α is a negative integer. It is the definition adopted by Hadamard [10]).

In particular, when $c = 0$, so that $f(z)$ is regular, we have

$$f_\alpha(z) = \sum \frac{\Gamma(n+1)}{\Gamma(n+1+\alpha)} a_n z^{n+\alpha}, \tag{3.13}$$

$$f_{-m}(z) = \sum_m^\infty n(n-1)\dots(n-m+1)\, a_n z^{n-m} = \left(\frac{d}{dz}\right)^m f(z). \tag{3.14}$$

It is with this case that we shall usually be concerned. It is then sometimes convenient to write

$$f_\alpha(z) = z^\alpha g_{(\alpha)}(z), \tag{3.15}$$

so that $g_{(\alpha)}(z)$ is regular.

We may extend the definitions of the means M_p and the classes L^p, given in § 1.2, to the functions $f_\alpha(z)$ and $g_{(\alpha)}(z)$. The means of f_α and $g_{(\alpha)}$ differ only by a factor r^α and are for our purposes equivalent.

If $\alpha > 0$ then

$$f_\alpha(z) = \frac{1}{\Gamma(\alpha)} \int_0^z (z-u)^{\alpha-1} f(u)\, du, \tag{3.16}$$

the path of integration being rectilinear and $(z-u)^{\alpha-1}$ having its principal value at the origin. This is the natural extension to analytic functions of Liouville's definition of the fractional integral.

---

[9]) There is no reason for this restriction except that real values of c are the only values relevant here.

[10]) Hadamard, 1, p. 56.

3.2. In 7 we used two definitions of the integral $F_\alpha(\theta)$, where $\alpha > 0$, of an integrable function $F(\theta)$, viz. Liouville's and Weyl's, which proved to be for our purposes equivalent. It is the second, which requires that $a_0 = 0$ or

$$\int_{-\pi}^{\pi} F(\theta)\, d\theta = 0,$$

which is relevant here. Then Weyl's definition is

$$F_\alpha(\theta) = \int_{-\infty}^{\theta} (\theta - u)^{\alpha-1} F(u)\, du,$$

and $F_\alpha(\theta)$ is integrable, periodic, and has mean-value 0. Further if, as here, $F(\theta)$ is of power series type, we have

$$i^\alpha F_\alpha(\theta) \sim \sum_1^\infty \frac{a_n}{n^\alpha} e^{ni\theta}\ ^{11});$$

and $i^\alpha F_\alpha(\theta)$ is, by Lemma A(3), the boundary function of the analytic function

$$\gamma_\alpha(z) = \sum_1^\infty \frac{a_n}{n^\alpha} z^n.$$

The factors i^α here, and z^α in (3.15), correspond to the difference between integration with respect to θ and with respect to $z = re^{i\theta}$.

The functions $g_{(\alpha)}(z)$ and $\gamma_\alpha(z)$ are (when, as here, $\alpha > 0$) equivalent for our present purposes. For if $M_1(f) \leqq C$ then

$$|a_n| \leqq \lim_{r\to 1} r^{-n} M_1(f) \leqq C;$$

and

$$\left| \frac{\Gamma(n+1)}{\Gamma(n+1+\alpha)} - \frac{1}{n^\alpha} \right| < \frac{K}{n^{1+\alpha}},$$

where $K = K(\alpha)$; so that

$$|g_{(\alpha)} - \gamma_\alpha| \leqq |a_0| + K \sum_1^\infty \frac{|a_n| r^n}{n^{1+\alpha}} \leqq KC.$$

We are thus able to translate many of the theorems of 7 into our present language.

Suppose in particular that $p > 1$ and

$$M_p(f) \leqq C.$$

---

[11]) See Weyl, 17. Weyl is concerned with general Fourier series.

Then $F(\theta)$ belongs to L^p and

$$\left(\frac{1}{2\pi}\int_{-\pi}^{\pi}|F(\theta)|^p\,d\theta\right)^{\frac{1}{p}}\leqq C.$$

If now $0<\alpha<\frac{1}{p}$ and

$$q=\frac{p}{1-p\alpha}>p,\quad \alpha=\frac{1}{p}-\frac{1}{q},$$

we have, by Theorem 4 of 7

$$\left(\frac{1}{2\pi}\int_{-\pi}^{\pi}|F_\alpha(\theta)|^q\,d\theta\right)^{\frac{1}{q}}\leqq K\left(\frac{1}{2\pi}\int_{-\pi}^{\pi}|F(\theta)|^p\,d\theta\right)^{\frac{1}{p}}\leqq KC,$$

where $K=K(p,\alpha)=K(p,q)$. It follows, by Lemma A (3), that

$$M_q(\gamma_\alpha)\leqq KC,$$

$$M_q(g_{(\alpha)})\leqq KC,$$

$$M_q(f_\alpha)=r^\alpha M_q(g_{(\alpha)})\leqq KC.$$

It is plain that a slight modification of the argument will prove that

$$M_q(f_\alpha)\leqq K M_p(f).$$

We thus obtain the following theorem.

Theorem 30. *If*

$$p>1,\quad 0<\alpha<\frac{1}{p},\quad q=\frac{p}{1-p\alpha},$$

and

$$M_p(f)\leqq C,$$

then

$$M_q(f_\alpha)\leqq K M_p(f)\leqq KC,$$

where $K=K(p,\alpha)$.

We shall return to this theorem later and extend it to all positive values of p. The case $p=2$, $q>2$ is required in the proof of Theorem 31.

4. The means M_p (continued).

4.1. Theorem 31. *Suppose that*

$$0<p<q,\quad \alpha=\frac{1}{p}-\frac{1}{q},\quad l\geqq p, \tag{4.11}$$

and

$$M_p(r)=M_p(r,f)\leqq C. \tag{4.12}$$

Then

$$\int_0^1 (M_q(r))^l (1-r)^{l\alpha-1}\,dr \leqq K C^l, \tag{4.13}$$

$$\int_0^1 (M(r))^l (1-r)^{\frac{l}{p}-1}\,dr \leqq K C^l. \tag{4.14}$$

In particular

$$\int_0^1 (M_q(r))^{\frac{pq}{q-p}}\,dr \leqq K C^{\frac{pq}{q-p}}, \tag{4.15}$$

$$\int_0^1 (M(r))^p\,dr \leqq K C^p, \tag{4.16}$$

$$\int_0^1 (M_q(r))^p (1-r)^{-\frac{p}{q}}\,dr \leqq K C^p, \tag{4.17}$$

$$\int_0^1 M_q(r)(1-r)^{\alpha-1}\,dr \leqq K C \qquad (p \leqq 1), \tag{4.18}$$

$$\int_0^1 M(r)(1-r)^{\frac{1}{p}-1}\,dr \leqq K C \qquad (p \leqq 1). \tag{4.19}$$

Here $K = K(p, q, l)$, $K = K(p, q)$, or $K = K(p)$, according to the parameters which appear in the conclusion.

The equations (4.14), (4.16), and (4.19) correspond to $q = \infty$.

4.2. (i) Whether q is finite or infinite, we can reduce the general case $l \geqq p$ to the special case $l = p$. For (supposing for example q finite) we have, by Theorem 27,

$$M_q(r) \leqq K C (1-r)^{-\frac{1}{p}+\frac{1}{q}}$$

and so

$$M_q^l(r)(1-r)^{l\alpha-1} \leqq K C^{l-p} M_q^p(r)(1-r)^{-\frac{p}{q}}.$$

This reduction is also valid for $q = \infty$.

(ii) If $f(z)$ has no zeros in the unit circle, we write

$$k = \frac{2q}{p} > 2, \qquad f = g^{\frac{2}{p}},$$

so that

$$M_2^2(g) = M_p^p(f) \leqq C^p,$$

and

$$\int_0^1 M_q^p(r, f)(1-r)^{-\frac{p}{q}}\,dr = \int_0^1 M_k^2(r, g)(1-r)^{-\frac{2}{k}}\,dr. \tag{4.21}$$

Suppose now that

$$g(z) = \sum c_n z^n = \sum \frac{\Gamma(n+1)}{\Gamma\left(n+\frac{3}{2}-\frac{1}{k}\right)} b_n z^n, \qquad h(z) = \sum b_n z^n,$$

so that

$$g(z) = z^{-\frac{1}{2}+\frac{1}{k}} h_{\frac{1}{2}-\frac{1}{k}}(z).$$

Taking $p = 2$, $q = k$ in Theorem 30, we find

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |g(r,\theta)|^k \, d\theta \leqq K\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |h(r,\theta)|^2 \, d\theta\right)^{\frac{1}{2}k},$$

or

$$M_k^k(r,g) \leqq K\left(\sum |b_n|^2 r^{2n}\right)^{\frac{1}{2}k}$$

Hence

$$\int_0^1 M_k^2(r,g)(1-r)^{-\frac{2}{k}} dr \leqq K \sum |b_n|^2 \int_0^1 r^{2n}(1-r)^{-\frac{2}{k}} dr$$

$$\leqq K \sum \left(\frac{\Gamma\left(n+\frac{3}{2}-\frac{1}{k}\right)}{\Gamma(n+1)}\right)^2 \frac{\Gamma(2n+1)}{\Gamma\left(2n+2-\frac{2}{k}\right)} |c_n|^2$$

$$\leqq K \sum |c_n|^2 = K M_2^2(1,g) \leqq K C^p;$$

and this and (4.21) give (4.17).

We can now extend the result to general f by Lemma B, writing $f = f_1 + f_2$, where

$$M_p(r,f_1) \leqq KC, \quad M_p(r,f_2) \leqq KC,$$

applying (4.17) to f_1 and f_2, and observing finally that

$$M_q^p(r,f) \leqq K M_q^p(r,f_1) + K M_q^p(r,f_2).$$

(iii) There remains the case $q = \infty$. If $f(z)$ has no zeros, we define $g(z)$ as before, when

$$M(f) = M^{\frac{2}{p}}(g),$$

$$\int_0^1 M^p(r,f)\, dr = \int_0^1 M^2(r,g)\, dr \leqq KC^p \ ^{12}).$$

The extension to the general case is effected as before.

---

[12]) In fact

$$\int_0^1 \mathfrak{G}^2(r)\, dr \leqq KC^p,$$

$\mathfrak{G}$ being the majorant $\sum |c_n| r^n$. See Theorem 15 of our paper 4. The result here is a direct corollary of Hilbert's double series theorem.

4.3. Before going further we require some additional preliminaries.

Theorem 32. (1) *If* $f(z)$ *is regular for* $r<1$, $p>0$,

$$M_p(f) \leqq C,$$

and

$$\mathfrak{F}(\theta) = \overline{\operatorname{bound}}_{r<1} |f(re^{i\theta})|,$$

then

$$\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \mathfrak{F}^p(\theta)\, d\theta\right)^{\frac{1}{p}} \leqq KC,$$

where $K = K(p)$.

The same result is true if $\mathfrak{F}(\theta)$ *is the upper bound of* $|f|$ *in the region* $S_\omega(\theta)$ *defined by* [13])

$$|\arg(e^{i\theta}-z)| \leqq \omega < \frac{1}{2}\pi, \qquad 0 \leqq |e^{i\theta}-z| \leqq 1,$$

but in this case with $K = K(p,\omega)$.

(2) *If*

$$\mathfrak{F}(r,\theta) = \operatorname*{Max}_{\varrho \leqq r} |f(\varrho e^{i\theta})|,$$

then

$$\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \mathfrak{F}^p(r,\theta)\, d\theta\right)^{\frac{1}{p}} \leqq KM_p(f).$$

(3) *When* $p>1$, *the corresponding results for harmonic and subharmonic functions are also true.*

Whe shall sometimes refer shortly to this theorem as "Max". It is merely a restatement of the principal function-theoretic theorems of our paper **9** (Theorems 17, 24, 25, 26, 27).

4.4. Suppose that $f(z)$ is regular for $r<1$, and that $\alpha>0$. We write

$$f_\alpha^*(z) = f_\alpha^*(re^{i\theta}) = f_\alpha^*(r,\theta) = \frac{1}{\Gamma(\alpha)}\int_0^r (r-\varrho)^{\alpha-1} |f(\varrho e^{i\theta})|\, d\varrho.$$

Plainly

$$|f_\alpha(z)| \leqq f_\alpha^*(z).$$

Similarly, if

$$u(z) = u(re^{i\theta}) = u(r,\theta) = \frac{1}{2}a_0 + \sum_1^\infty (a_n \cos n\theta + b_n \sin n\theta) r^n$$

---

[13]) $S_\omega(\theta)$ is a kite-shaped region of fixed size and shape pointing inwards from the point $e^{i\theta}$ and with a boundary which does not touch the unit circle. It diminishes when ω decreases, and it reduces to the radius vector when $\omega = 0$.

is a (real or complex) harmonic function regular for $r<1$, we write

$$u_\alpha(z)=u_\alpha(r,\theta)=\frac{1}{\Gamma(\alpha)}\int_0^r(r-\varrho)^{\alpha-1}u(\varrho,\theta)\,d\varrho,$$

$$u_\alpha^*(z)=u_\alpha^*(r,\theta)=\frac{1}{\Gamma(\alpha)}\int_0^r(r-\varrho)^{\alpha-1}|u(\varrho,\theta)|\,d\varrho,$$

and again

$$|u_\alpha(z)|\leqq u_\alpha^*(z).$$

We define the means

$$M_p(r,f_\alpha^*),\quad M_p(r,u_\alpha),\quad M_p(r,u_\alpha^*)$$

in a way similar to that in which we defined $M_p(r,f)$ and $M_p(r,u)$ in § 1.2.

4.5. We come now to the principal theorems of the paper.

Theorem 33. *If*

$$p>0,\quad 0<\alpha<\frac{1}{p},\quad q=\frac{p}{1-p\alpha},$$

and

$$M_p(f)\leqq C,$$

then

$$M_q(f_\alpha)\leqq KC,$$

where

$$K=K(p,\alpha)=K(p,q).$$

Theorem 34. *Under the same conditions*

$$M_q(f_\alpha^*)\leqq KC.$$

Theorem 35. *If*

$$p>1,\quad 0<\alpha<\frac{1}{p},\quad q=\frac{p}{1-p\alpha},$$

and

$$M_p(u)\leqq C,$$

then

$$M_q(u_\alpha)\leqq KC.$$

Theorem 36. *Under the conditions of Theorem* 35

$$M_q(u_\alpha^*)\leqq KC.$$

We shall sometimes refer to Theorems 33 and 34 as

$$\text{“}p\to q\text{”},\quad \text{“}p\overset{*}{\to}q\text{”}$$

respectively. It is plain that Theorem 34 is stronger than and includes Theorem 33, and that, when $p>1$, Theorems 35 and 36 include Theorems 33 and 34 respectively.

Proof of Theorems 33—36.

4.61. (1) *The case* $1 < p < q$. When $p > 1$, all the theorems are easy deductions from Theorem 30 (and so of Theorem 4 of 7).

In the first place, Theorem 33 reduces to Theorem 30.

Next, suppose that u is harmonic and real and that $M_p(u) \leqq C$, and let v be the conjugate of u which vanishes at the origin, and $f = u + iv$. By M. Riesz's theorem,

$$M_p(f) \leqq KC,$$

where $K = K(p)$, and so, by Theorem 33,

$$M_p(f_\alpha) \leqq KC.$$

Since

$$f_\alpha(r,\theta) = \frac{1}{\Gamma(\alpha)} \int_0^r (r-\varrho)^{\alpha-1} (u(\varrho,\theta) + iv(\varrho,\theta))\, d\varrho = u_\alpha(r,\theta) + iv_\alpha(r,\theta),$$

we have

$$|u_\alpha| \leqq |f_\alpha|,$$

$$M_p(u_\alpha) \leqq M_p(f_\alpha) \leqq KC.$$

This proves Theorem 35 when u is real, and the general form of the theorem follows by considering the real and imaginary parts separately.

Next, suppose that $M_p(u) \leqq C$, and that $U(\theta) = u(e^{i\theta})$ is the boundary function of u. If $W(r,\theta)$ is the harmonic function which has the boundary values $|U(\theta)|$, then

$$M_p(W) \leqq C.$$

Also $|u(r,\theta)|$ is subharmonic, and therefore, by F. Riesz's fundamental theorem on subharmonic functions [14]),

$$|u(r,\theta)| \leqq W(r,\theta);$$

and therefore

$$u_\alpha^*(r,\theta) = \frac{1}{\Gamma(\alpha)} \int_0^r (r-\varrho)^{\alpha-1} |u(\varrho,\theta)|\, d\varrho \leqq W_\alpha(r,\theta),$$

$$M_q(u_\alpha^*) \leqq M_q(W_\alpha) \leqq KC,$$

by Theorem 35. This proves Theorem 36 (and so Theorem 34, when $p > 1$).

Alternatively, we may argue as follows. Assuming u to be real, we have, by Lemma A (6), $u = u_1 - u_2$, where u_1 and u_2 are harmonic and positive and

$$M_p(u_1) \leqq C, \quad M_p(u_2) \leqq C.$$

---

[14]) F. Riesz, 14.

Then

$$u_\alpha^* \leqq (u_1)_\alpha + (u_2)_\alpha,$$

$$M_q(u_\alpha^*) \leqq M_q((u_1)_\alpha) + M_q((u_2)_\alpha) \leqq KC,$$

by Theorem 35.

This argument is more elementary, since it avoids any appeal to the theory of sub-harmonic functions; but we shall have to refer to the first argument later.

4.62. (2) *The case* $p \leqq 1 < q$. By the integral form of Minkowski's inequality, we have

$$\begin{aligned}
M_q(f_\alpha^*) &= \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} f_\alpha^{*q}\, d\theta\right)^{\frac{1}{q}} \\
&= \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} d\theta \left(\frac{1}{\Gamma(\alpha)}\int_0^r (r-\varrho)^{\alpha-1} |f(\varrho e^{i\theta})|\, d\varrho\right)^q\right)^{\frac{1}{q}} \\
&\leqq \frac{1}{\Gamma(\alpha)}\int_0^r (r-\varrho)^{\alpha-1}\, d\varrho \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(\varrho e^{i\theta})|^q\, d\theta\right)^{\frac{1}{q}} \\
&= \frac{1}{\Gamma(\alpha)}\int_0^r (r-\varrho)^{\alpha-1} M_q(\varrho, f)\, d\varrho \\
&= \frac{r^\alpha}{\Gamma(\alpha)}\int_0^1 (1-t)^{\alpha-1} M_q(rt, f)\, dt \\
&\leqq \frac{1}{\Gamma(\alpha)}\int_0^1 (1-t)^{\alpha-1} M_q(t, f)\, dt \leqq KC,
\end{aligned}$$

by Theorem 31 (4.18).

It should be observed that the proof of Theorem 31 depends upon Theorem 30 (*i. e.* upon the case $p > 1$ of Theorem 33).

4.63. (3) *The case* $p < q \leqq 1$. We give two proofs, each of which has interesting features.

(i) Suppose that $0 < r < 1$ and

$$r_n = (1 - 2^{-n})\, r \qquad (n = 0, 1, 2, \ldots),$$

so that $r_0 = 0$ and $r_n \to r$ when $n \to \infty$. Then

$$r - r_n = r_n - r_{n-1} = 2^{-n} r,$$

so that the ratio of two factors of the form $(r - r_\nu)^c$, where ν is $n-1$, n, or $n+1$, lies between positive bounds depending on c only.

In these circumstances we have

$$f_\alpha^*(r,\theta) = K\int_0^r (r-\varrho)^{\alpha-1}|f(\varrho,\theta)|\,d\varrho$$

$$= K\sum_1^\infty \int_{r_{n-1}}^{r_n} (r-\varrho)^{\alpha-1}|f(\varrho,\theta)|\,d\varrho = K\sum_1^\infty J_n,$$

say. Also

$$J_n \leqq K(r_n - r_{n-1})(r-r_n)^{\alpha-1}\mu_n \leqq K2^{-n\alpha}r^\alpha\mu_n,$$

where

$$\mu_n = \mu_n(\theta) = \operatorname*{Max}_{r_{n-1}\leqq\varrho\leqq r_n}|f(\varrho,\theta)|;$$

and so (since $q \leqq 1$)

$$f_\alpha^{*q} \leqq K\sum J_n^q \leqq K\sum 2^{-nq\alpha}r^{q\alpha}\mu_n^q,$$

$$M_q^q(f_\alpha^*) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f_\alpha^{*q}\,d\theta \leqq K\sum_1^\infty 2^{-nq\alpha}r^{q\alpha}\frac{1}{2\pi}\int_{-\pi}^{\pi}\mu_n^q\,d\theta$$

$$\leqq K\sum_1^\infty 2^{-nq\alpha}r^{q\alpha}M_q^q(r_n,f),$$

by "Max" (Theorem 32). Hence

$$M_q^q(f_\alpha^*) \leqq K\sum_1^\infty (r_{n+1}-r_n)(r-r_{n+1})^{q\alpha-1}M_q^q(r_n,f)$$

$$\leqq K\sum_0^\infty \int_{r_n}^{r_{n+1}} (r-\varrho)^{q\alpha-1}M_q^q(\varrho,f)\,d\varrho$$

$$= K\int_0^r (r-\varrho)^{q\alpha-1}M_q^q(\varrho)\,d\varrho = Kr^{q\alpha}\int_0^1 (1-t)^{q\alpha-1}M_q^q(rt)\,dt$$

$$\leqq K\int_0^1 (1-t)^{q\alpha-1}M_q^q(t)\,dt \leqq KC^q,$$

by Theorem 31 (4.13).

The proof as stated depends on the deep theorem "Max", and it may be interesting to notice that a logical simplification is possible. We may use, instead of Theorem 32, the theorem:

If

$$r_2 < r_1 < r < 1$$

and

$$0 < A_1 < \frac{r-r_1}{r_1-r_2} < A_2,$$

where the A are absolute constants; and if

$$\nu = \nu(\theta) = \operatorname*{Max}_{r_2 \leqq \varrho \leqq r_1} |f(\varrho, \theta)|,$$

then

$$\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \nu^p \, d\theta\right)^{\frac{1}{p}} \leqq K M_p(r, f).$$

This theorem is included in Theorem 32 but is considerably easier, since the proof does not depend upon the maximal principle involved in that of Theorem 32.

To apply it for our present purpose, we take r_2, r_1, r to be r_{n-1}, r_n, r_{n+1}, when we obtain

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \mu_n^q \, d\theta \leqq K M_q^q(r_{n+1}, f).$$

The changes in the subsequent argument owing to the occurrence of $n+1$ in place of n are trivial.

4.64. (ii) We may suppose that $f(z)$ has no zeros in the unit circle, the general case being plainly reducible to this one by Lemma B.

We can choose s so that

$$s > 1, \quad \frac{q}{q-p+pq} < s < \frac{q}{q-p}.$$

Then, if

$$\tau = \frac{sp}{sq-q},$$

we have

$$\tau > 1,$$

$$q\tau' = \frac{q\tau}{\tau-1} = \frac{spq}{sp-sq+q} > 1$$

and

$$\frac{p}{q\tau} + \frac{1}{s} = \frac{s-1}{s} + \frac{1}{s} = 1$$

If

$$f = g^s$$

then

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |g|^{sp} \, d\theta \leqq C^p.$$

Also

$$f_\alpha^* = K\int_0^r |g(\varrho e^{i\theta})|^s (r-\varrho)^{\alpha-1} \, d\varrho$$

$$\leqq K\mathfrak{G}^{s-1}\int_0^r |g(\varrho e^{i\theta})| \, (r-\varrho)^{\alpha-1} \, d\varrho,$$

where

$$\mathfrak{G} = \mathfrak{G}(r, \theta) = \max_{\varrho \leqq r} |g(\varrho e^{i\theta})|;$$

so that

$$f_\alpha^{*q} \leqq K \mathfrak{G}^{(s-1)q} g_\alpha^{*q}.$$

Hence, by Hölder's inequality,

$$\int f_\alpha^{*q} d\theta \leqq K \left(\int \mathfrak{G}^{(s-1)q\tau} d\theta \right)^{\frac{1}{\tau}} \left(\int g_\alpha^{*q\tau'} d\theta \right)^{\frac{1}{\tau'}} \tag{4.641}$$

$$= K \left(\int \mathfrak{G}^{sp} d\theta \right)^{\frac{1}{\tau}} \left(\int g_\alpha^{*q\tau'} d\theta \right)^{\frac{1}{\tau'}}.$$

Now

$$\left(\int \mathfrak{G}^{sp} d\theta \right)^{\frac{1}{\tau}} \leqq K \left(\int |g|^{sp} d\theta \right)^{\frac{1}{\tau}} \leqq K C^{\frac{p}{\tau}}, \tag{4.642}$$

by Theorem 32. Also

$$q\tau' = \frac{spq}{sp - sq + q} = \frac{sp}{1 - sp\alpha}.$$

Hence[15])

$$\left(\int g_\alpha^{*q\tau'} d\theta \right)^{\frac{1}{\tau'}} \leqq K \left(\int |g|^{sp} d\theta \right)^{\frac{q}{sp}} \leqq K C^{\frac{q}{s}}. \tag{4.643}$$

From (4.641), (4.642) and (4.643) we deduce

$$M_q(f_\alpha^*) \leqq K \left(K C^{\frac{p}{\tau}} \right)^{\frac{1}{q}} (KC)^{\frac{1}{s}} = KC;$$

which proves the theorem.

Additional remarks on the propositions $p \to q$ and $p \overset{*}{\to} q$.

4.7. (1) It follows immediately from the definitions that $p \to q$ is transitive, *i. e.* that $p \to q$ and $q \to r$ imply $p \to r$, while this relation is not immediate for $p \overset{*}{\to} q$. On the other hand we can show very easily (without being able to prove the proposition itself) that the latter proposition is convex in q, *i. e.* that $p \overset{*}{\to} q_1$ and $p \overset{*}{\to} q_2$, where $p < q_1 < q_2$, imply $p \overset{*}{\to} q$ for $q_1 < q < q_2$. In fact if

$$q_1 = \frac{p}{1 - p\alpha_1}, \quad q_2 = \frac{p}{1 - p\alpha_2}, \quad \alpha_1 = \frac{1}{p} - \frac{1}{q_1}, \quad \alpha_2 = \frac{1}{p} - \frac{1}{q_2},$$

and

$$\alpha = \delta\alpha_1 + (1 - \delta)\alpha_2,$$

---

[15]) By one or other of the cases of the theorem proved already; the case $1 < p < q$ or $p \leqq 1 < q$ according as $sp > 1$ or $sp \leqq 1$.

where $0 < \delta < 1$, so that

$$\frac{1}{q} = \frac{\delta}{q_1} + \frac{1-\delta}{q_2},$$

we have at once, from Hölder's inequality,

$$f_\alpha^* = K \int_0^r (r-\varrho)^{\alpha-1} |f(\varrho e^{i\theta})| \, d\varrho \leqq K f_{\alpha_1}^{*\delta} f_{\alpha_2}^{*1-\delta},$$

$$\int f_\alpha^{*q} \, d\theta \leqq K \left(\int f_{\alpha_1}^{*q_1} \, d\theta \right)^{\frac{\delta q}{q_1}} \left(\int f_{\alpha_2}^{*q_2} \, d\theta \right)^{\frac{(1-\delta)q}{q_2}} \leqq K \left(\int |f|^p \, d\theta \right)^{\frac{q}{p}}.$$

There is no such immediate proof of the convexity of $p \to q$.

(2) The proof that $p \to q$ is true generally depends on a rather formidable chain of theorems. There are simpler proofs of a number of special cases of the theorem which are of some intrinsic interest, and it may be worth while to notice one or two of them explicitly. In the remarks which follow we suppose, for simplicity of writing, that $a_0 = 0$ and that our series begin from $n = 1$, and we state our theorems as convergence theorems instead of as inequalities.

(i) The theorem $1 \to 2$ asserts that

$$\sum \frac{|a_n|^2}{n} < \infty$$

for any $f(z)$ of L. Since $a_n = o(1)$, this is included in the stronger theorem (Theorem 16 of our paper 4 [16]), with $\lambda = 1$) that

$$\sum \frac{|a_n|}{n} < \infty. \tag{4.71}$$

(ii) Similarly the theorem $p \to 2$, where $p < 1$, asserts that

$$\sum n^{-\frac{2}{p}+1} |a_n|^2 < \infty$$

for any $f(z)$ of L^p. Since, by Theorem 28, $a_n = o\left(n^{\frac{1}{p}-1}\right)$, this is included in the stronger theorem (Theorem 16 of 4) that

$$\sum n^{p-2} |a_n|^p < \infty.$$

An intermediate theorem is that

$$\sum n^{-\frac{1}{p}} |a_n| < \infty. \tag{4.72}$$

We stated this in our original note 3 on fractional integrals, but do not seem to have ever published a proof.

---

[16]) Some simplifications of the argument of 4 are indicated in 6.

(iii) We can deduce the theorem $p \overset{*}{\rightarrow} 1$, where $p < 1$, from $p \rightarrow 2$, as follows. We have to prove that

$$\int f^*_{\frac{1}{p}-1}\, d\theta,$$

where

$$f^*_{\frac{1}{p}-1} = K \int_0^r (r-\varrho)^{\frac{1}{p}-2} |f(\varrho e^{i\theta})|\, d\varrho,$$

is bounded for any $f(z)$ of L^p. It will be sufficient, after Lemma B, to prove the result on the assumption that $f(z)$ has no zeros in the unit circle. Writing then

$$f = \varphi^2, \qquad \varphi = \sum c_n z^n,$$

so that φ is L^{2p}, we have

$$\begin{aligned}
\int f^*_{\frac{1}{p}-1}\, d\theta &= K \int_0^r (r-\varrho)^{\frac{1}{p}-2} d\varrho \int |\varphi(\varrho e^{i\theta})|^2 d\theta \\
&= K \int_0^r (r-\varrho)^{\frac{1}{p}-2} \Big(\sum |c_n|^2 \varrho^{2n} \Big) d\varrho \\
&= K \sum \frac{\Gamma\left(\frac{1}{p}-1\right)\Gamma(2n+1)}{\Gamma\left(2n+\frac{1}{p}\right)} |c_n|^2 r^{2n+\frac{1}{p}-1} \\
&\leqq K \sum n^{1-\frac{1}{p}} |c_n|^2 r^{2n} \leqq K \int \Big| \varphi_{\frac{1}{2p}-\frac{1}{2}} \Big|^2 d\theta < K,
\end{aligned}$$

by $2p \rightarrow 2$.

(iv) If $p < 1$ and f is L^p, and we assume $p \rightarrow 1$, then $f_{\frac{1}{p}-1}$ is L. The convergence of the series (4.72) then follows from (4.71). This is however not an economical line of proof.

(3) We can if we please prove a still stronger proposition than $p \overset{*}{\rightarrow} q$, which we may denote by $p \overset{**}{\rightarrow} q$; viz.

Theorem 37. *If the conditions of Theorem* 33 *are satisfied then*

$$M_q(f^{**}_\alpha) \leqq KC,$$

where

$$f^{**}_\alpha = \frac{1}{\Gamma(\alpha)} \int_0^r (r-\varrho)^{\alpha-1} \operatorname*{Max}_{t \leqq \varrho} |f(te^{i\theta})|\, d\varrho.$$

We shall be content to sketch the proof of this theorem, which depends on a lemma of more interest than the theorem itself.

Lemma C. *If* $\lambda > 0$, *and*

$$\mathfrak{F} = \mathfrak{F}(r, \theta) = \operatorname*{Max}_{\varrho \leqq r} |f(\varrho e^{i\theta})|$$

is the function of Theorem 32, *then* $\mathfrak{F}^\lambda$ *is sub-harmonic.*

We suppose that P is the point $re^{i\theta}$, and write

$$\mathfrak{F}(P) = \mathfrak{F}(re^{i\theta}) = |f(te^{i\theta})|,$$

where $t = t(r, \theta)$ gives the maximum of $|f(te^{i\theta})|$. We have to show that

$$\mathfrak{F}_P^\lambda \leqq \operatorname*{Av}_{\Gamma} \mathfrak{F}_{P'}^\lambda, \tag{4.73}$$

where "Av." denotes an average over points P' of a sufficiently small circle Γ with centre P.

There are three possibilities as regards t, viz. (a) $t = 0$, (b) $t = r$, and (c) $0 < t < r$. Since $\mathfrak{F} \geqq |f(0)|$, (4.73) is certainly true in case (a). In case (b) $\mathfrak{F}(P) = |f(P)|$, and, since $|f|^\lambda$ is sub-harmonic,

$$\mathfrak{F}^\lambda(P) = |f(P)|^\lambda \leqq \operatorname*{Av}_{\Gamma} |f(P')|^\lambda \leqq \operatorname*{Av}_{\Gamma} \mathfrak{F}^\lambda(P').$$

Finally, in case (c), we denote the point $te^{i\theta}$ by Q and construct the circle γ which has its centre at Q and is in similitude with Γ about the origin. If Q' is the point of γ in similitude with P', then $|f(Q')| \leqq \mathfrak{F}(P')$, and so, since corresponding arcs of γ and Γ are proportional,

$$\mathfrak{F}^\lambda(P) = |f(Q)|^\lambda \leqq \operatorname*{Av}_{\gamma} |f(Q')|^\lambda \leqq \operatorname*{Av}_{\Gamma} \mathfrak{F}^\lambda(P').$$

It may be verified by direct calculation that, in case (c),

$$\nabla^2 \mathfrak{F}^\lambda = \frac{t^2}{r^2} \nabla^2 |f(te^{i\theta})|^\lambda + \frac{\lambda |f|^{\lambda-2}}{r^2} \frac{K^2 + L^2}{K},$$

where ∇^2 is Laplace's operator,

$$K = -\frac{1}{2} t^2 \frac{\partial^2}{\partial t^2} |f(te^{i\theta})|^2 \geqq 0, \qquad L = u \frac{\partial^2 v}{\partial \theta^2} - v \frac{\partial^2 u}{\partial \theta^2},$$

and u and v are the components of $f(te^{i\theta})$.

It follows from the lemma that, if u^* is the harmonic function which has the boundary values $\mathfrak{F}(1, \theta)$, then $\mathfrak{F} \leqq u^*$ inside the unit circle. The proof of Theorem 37 for $p > 1$, and also the subsequent stages of the proof, then run parallel to the corresponding stages of the proof of Theorem 34, if we appeal to Theorem 32 at the appropriate places.

An extension of Theorem 33.

4.8. Theorem 38. *If*

$$p > 0, \quad \alpha > 0, \quad 0 < \alpha < \frac{1}{p}, \quad q = \frac{p}{1 - p\alpha}$$

and

$$M_p(f) \leqq C(1 - r)^{-\alpha}, \tag{4.81}$$

then

$$M_q(f_\alpha) \leqq K C (1-r)^{-\alpha}, \tag{4.82}$$

where

$$K = K(p, \alpha, \alpha), = K(p, \alpha, q).$$

This is a corollary of Theorem 33. We write

$$\varrho = \frac{1}{2}(1+r), \qquad \varphi(z) = f(\varrho z).$$

Then

$$M_p(f) \leqq C(1-\varrho)^{-\alpha}$$

for $r < \varrho$, or

$$M_p(\varphi) \leqq C(1-\varrho)^{-\alpha}$$

for $r < 1$; so that, by Theorem 33

$$M_q(\varphi_\alpha) \leqq K C (1-\varrho)^{-\alpha}$$

for $r < 1$. But

$$f_\alpha(\varrho z) = \varrho^\alpha \varphi_\alpha(z)$$

and so

$$M_q(f_\alpha(\varrho z)) \leqq K C (1-\varrho)^{-\alpha} \leqq K C (1-\varrho r)^{-\alpha}$$

for $r < 1$. This is equivalent to (4.82).

5. The order of the integrals and derivatives of $f(z)$.

5.1. The theorems proved in this section are not difficult, but they require some preliminaries.

We write

$$f^\beta(z) = f_{-\beta}(z),$$

so that

$$f^\beta(z) = \left(\frac{d}{dz}\right)^\beta f(z)$$

when β is a positive integer.

Suppose that $\beta > 0$. Then

$$f^\beta(z) = \sum \frac{\Gamma(n+1)}{\Gamma(n+1-\beta)} a_n z^{n-\beta}.$$

But

$$\frac{\Gamma(n+1)}{\Gamma(n+1-\beta)} = \frac{\Gamma(1+\beta)}{2\pi i} \int_C u^n (u-1)^{-1-\beta}\, du,$$

where C is a loop from $u = 0$ round $u = 1$ in the positive direction, and $(u-1)^{-1-\beta}$ has its principal value at the positive point of C. Hence, provided $|zu| < 1$ at all points of C (a condition we can certainly secure

if $|z|<1$) we have

$$f^{\beta}(z)=\frac{\Gamma(1+\beta)}{2\pi i}z^{-\beta}\int_{C}f(zu)(u-1)^{-1-\beta}du$$

$$=\frac{\Gamma(1+\beta)}{2\pi i}\int_{C_z}f(w)(w-z)^{-1-\beta}dw,$$

where C_z is a loop from $w=0$ round $w=z$.

When β is an integer, we can replace C_z by a circle (so obtaining the ordinary formula for $f^{\beta}(z)$). In general we may replace it by the contour Γ_z formed by the circle

$$|w|=\varrho \qquad (r<\varrho<1)$$

and a double lacet from $w=0$ to the point on this circle opposite to z, viz. $-\varrho e^{i\theta}$. We call the circle and the two edges of the lacet γ_z, $\Gamma_{1,z}$, $\Gamma_{2,z}$.

Theorem 39. *If* $a>0$, $a+\beta>0$, *then the hypotheses*

(5.11) $$f(z)=O((1-r)^{-a})$$

and

(5.12.) $$f^{\beta}(z)=O((1-r)^{-a-\beta})$$

are equivalent. If $a=0$, $a+\beta>0$ *then* (5.11) *implies* (5.12); *and if* $a>0$, $a+\beta=0$ *then* (5.12) *implies* (5.11); *the converse implication being in either case untrue.*

(i) Suppose that $a\geqq 0$, $\beta>0$, and that (5.11) is true; and take

$$\varrho=\tfrac{1}{2}(1+r).$$

Then $f^{\beta}(z)$ is a constant multiple of

$$\int_{\Gamma_z}\frac{f(w)}{(w-z)^{1+\beta}}dw=\int_{\gamma_z}+\int_{\Gamma_{1,z}}+\int_{\Gamma_{2,z}}.$$

The two last integrals are

$$O\left(\int_0^{\varrho}(1-t)^{-a}dt\right),$$

and so negligible. The first is

$$O\left\{(1-r)^{-a}\int_{-\pi+\theta}^{\pi+\theta}\frac{d\varphi}{|\varrho e^{i\varphi}-re^{i\theta}|^{1+\beta}}\right\}$$

$$=O((1-r)^{-a}(\varrho-r)^{-\beta})=O((1-r)^{-a-\beta}).$$

(ii) Suppose that $a>0$, $\beta>0$, and that (5.12) is true, and write

$$f^{\beta}=g, \qquad f=g_{\beta}.$$

Then

$$f(z)=\frac{1}{\Gamma(\beta)}\int_0^z g(w)(z-w)^{\beta-1}dw=O\Big(\int_0^r(1-t)^{-\alpha-\beta}(r-t)^{\beta-1}dt\Big).$$

If we write $x=\frac{1-t}{1-r}$, the last integral becomes

$$(1-r)^{-\alpha}\int_1^{\frac{1}{1-r}}x^{-\alpha-\beta}(x-1)^{\beta-1}dx<(1-r)^{-\alpha}\int_1^{\infty}=K(1-r)^{-\alpha}.$$

Hence

$$f(z)=O((1-r)^{-\alpha}).$$

A moment's consideration shows that (i) and (ii) cover all the positive assertions of the theorem. The negative assertion is established by the trivial example

$$f=\log\frac{1}{1-z},\quad f'=\frac{1}{1-z}.$$

It may be worth observing (in view of our later Theorem 46) that $f(z)=O(1)$ does not necessarily imply $f^{\beta}(z)=o((1-r)^{-\beta})$ for $\beta>0$: consider, for example, the function

$$f(z)=(1-z)^{i}.$$

Lipschitz conditions.

5.2. Suppose that $0\leqq k\leqq 1$. We say that $f(z)$ satisfies a Lipschitz condition of order k in the unit circle, or belongs to the complex class $\mathrm{Lip}\,k$, if

$$|f(z')-f(z)|\leqq K|z'-z|^{k}, \tag{5.21}$$

where K is independent of z and z', for $|z|\leqq 1$, $|z'|\leqq 1$. If

$$f(z')-f(z)=o(|z'-z|^{k}),$$

when $z'\to z$, uniformly in z, we say that $f(z)$ belongs to $\mathrm{Lip}^{*}k$. The classes $\mathrm{Lip}\,0$ and $\mathrm{Lip}^{*}0$ are those of bounded and continuous functions. In the first case we assert (5.21) in the first instance only for $|z|<1$, $|z'|<1$; it is then satisfied automatically for those z and z' on the circle for which the boundary function exists.

Theorem 40. *In order that $F(\theta)=f(e^{i\theta})$ should belong to the real class* $\mathrm{Lip}\,k$, *where $0<k\leqq 1$, it is necessary and sufficient that*

$$f'(z)=O((1-r)^{k-1}). \tag{5.22}$$

This is Theorem 4 of 8. As we did not prove it explicitly there, and the proof is short, we give it here.

The real class Lip 1 is the same as the class of integrals of bounded functions. Also, if $f(e^{i\theta})$ belongs to the real class Lip k, it follows from the principle of the maximum modulus, applied to the function

$$\varphi(z) = f(ze^{ih}) - f(z),$$

that

$$|f(re^{i(\theta+h)}) - f(re^{i\theta})| \leqq Kh^k \tag{5.21}$$

uniformly for $0 \leqq r < r + h < 1$ and all θ, *i. e.* that $f(re^{i\theta})$, considered as a function of θ, belongs to Lip k, uniformly in r. If in particular $k = 1$, $f'(re^{i\theta})$ is uniformly bounded. This proves the theorem when $k = 1$. We may therefore suppose that $k < 1$.

(i) Suppose that $F(\theta)$ belongs to Lip k. We have, as on p. 626 of 8 [17]),

$$f'(re^{i\theta}) = \frac{e^{-i\theta}}{2\pi} \int_{-\pi}^{\pi} \frac{e^{i\psi}}{(e^{i\psi} - r)^2} (F(\theta + \psi) - F(\theta))\, d\psi = O\left(\int_{-\pi}^{\pi} \frac{|\psi|^k d\psi}{|e^{i\psi} - r|^2} \right)$$

$$= O((1 - r)^{k-1}).$$

(ii) Suppose that (5.22) is true. It is plain that $f(z)$ tends uniformly to a limit when $r \to 1$, and that $F(\theta)$ is continuous. Also, if $h > 0$,

$$F(\theta + h) - F(\theta) = \int_{e^{i\theta}}^{e^{i(\theta+h)}} f'(z)\, dz = \left(\int_{(1)} + \int_{(2)} + \int_{(3)} \right) f'(z)\, dz = I_1 + I_2 + I_3;$$

the paths of integration in I_1, I_2 and I_3 being respectively (1) the radius from $e^{i\theta}$ to $(1 - h)e^{i\theta}$, (2) the circle $r = 1 - h$ from $(1 - h)e^{i\theta}$ to $(1 - h)e^{i(\theta+h)}$, and (3) the radius from $(1 - h)e^{i(\theta+h)}$ to $e^{i(\theta+h)}$. It is plain that

$$I_2 = O(h \cdot h^{k-1}) = O(h^k);$$

and I_1 and I_3 are each

$$O\left(\int_{1-h}^{1} (1 - r)^{k-1}\, dr \right) = O(h^k).$$

5.3. Theorem 41. *If $F(\theta) = f(e^{i\theta})$ belongs to the real class* Lip k, *then $f(z)$ belongs to the complex class* Lip k.

We may suppose $k > 0$, the result being classical when $k = 0$. It is plainly enough to prove (5.21) and

$$|f((r + h)e^{i\theta}) - f(re^{i\theta})| \leqq Kh^k, \tag{5.31}$$

uniformly for $0 \leqq r < r + h < 1$ and all θ.

---

[17]) The notation is now different.

In proving (5.31), we may take $\theta = 0$ without real loss of generality. We have then, by Theorem 40,

$$|f(r+h) - f(r)| \leqq \int_r^{r+h} |f'(t)|\, dt \leqq K \int_r^{r+h} (1-t)^{k-1}\, dt.$$

If $r+h \leqq \frac{1}{2}(1+r)$, then $h \leqq \frac{1}{2}(1-r)$ and $1-t \geqq \frac{1}{2}(1-r)$; and

$$\int_r^{r+h} (1-t)^{k-1}\, dt < Kh(1-r)^{k-1} < Kh \cdot h^{k-1} = Kh^k.$$

If $r+h \geqq \frac{1}{2}(1+r)$ then $h \geqq \frac{1}{2}(1-r)$ and

$$\int_r^{r+h} (1-t)^{k-1}\, dt < K(1-r)^k < Kh^k.$$

Hence (5.31) is true in either case.

5.4. Theorem 42. *If* $0 < k \leqq 1$ *and*

$$f(re^{i\theta}) - f(r'e^{i\theta}) = O((r-r')^k),$$

uniformly for $0 \leqq r' < r \leqq 1$ *and all* θ, *then* $f(z)$ *belongs to* Lip k.

In other words, if $f(z)$ satisfies a Lipschitz condition uniformly on every radius, it satisfies one in the circle. We do not actually need this, theorem but it has a certain interest and is in a sense a partial converse of Theorem 41.

When $k = 1$, the hypothesis and the conclusion are each equivalent to the assertion that $f'(z)$ is bounded. We may therefore suppose $k < 1$.

If $0 < h < 1$, we have

$$f(re^{i\theta}) - f((r-hr)\, e^{i\theta}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(f(e^{i\varphi}) - f((1-h)\, e^{i\varphi})\right) P(r, \varphi - \theta)\, d\varphi.$$

Differentiating with respect to θ, we obtain

$$rf'(re^{i\theta}) - (r-hr) f'((r-hr)\, e^{i\theta})$$

$$= \frac{e^{-i\theta}}{2\pi i} \int_{-\pi}^{\pi} \left(f(e^{i\varphi}) - f((1-h)\, e^{i\varphi})\right) \frac{\partial P}{\partial \theta}\, d\varphi.$$

But

$$\left|\frac{\partial P}{\partial \theta}\right| = \left|\frac{2r(1-r^2)\sin(\varphi-\theta)}{(1-2r\cos(\varphi-\theta)+r^2)^2}\right| \leqq \frac{AP}{1-r},$$

where A is an absolute constant. Hence

$$(5.41) \qquad rf'(re^{i\theta}) - (r-hr) f'((r-hr)\, e^{i\theta}) = O\left(\frac{h^k}{1-r}\right).$$

We now take

$$r_n = 1 - 2^{-n}, \quad r_n(1-h_n) = r_{n-1},$$

so that

$$h_n = \frac{2^{-n}}{1-2^{-n}} = O(2^{-n});$$

and write

$$r = r_n, \quad h = h_n \qquad (n = 2, 3, \ldots, N)$$

in (5.41). Adding the resulting equations, we obtain

$$\begin{aligned} r_N f'(r_N e^{i\theta}) &= O(2^{N(1-k)} + 2^{(N-1)(1-k)} + \ldots) + O(1) \\ &= O(2^{N(1-k)}) = O((1-r_N)^{k-1}). \end{aligned}$$

If we observe that $1-r$ varies only by a factor 2 in the interval (r_{N-1}, r_N), and use the principle of the maximum modulus, we obtain

$$f'(z) = O((1-r)^{k-1}),$$

and the conclusion then follows from Theorem 40.

5.5. Theorem 43. *If* $0 < k < 1$, $\beta > k$, *then a necessary and sufficient condition that* $f(z)$ *should belong to* Lip k *is that*

$$f^{\beta}(z) = O((1-r)^{k-\beta}).$$

After Theorem 39, we may suppose that $\beta = 1$, and the conclusion then follows from Theorems 40 and 41.

Theorem 44. *If* $0 < k < 1$, $0 < k - \beta < 1$, *then a necessary and sufficient condition that* $f(z)$ *should belong to* Lip k *is that* $f^{\beta}(z)$ *should belong to* Lip $(k-\beta)$.

Suppose, for example, that $\beta > 0$. In order that $f(z)$ should belong to Lip k, it is necessary and sufficient (by Theorems 40 and 41) that

$$f'(z) = O((1-r)^{k-1});$$

and therefore (by Theorem 39) that

$$f^{\beta+1}(z) = O((1-r)^{k-\beta-1});$$

and therefore (by Theorems 40 and 41) that $f^{\beta}(z)$ should belong to Lip $(k-\beta)$.

There are partial implications corresponding to the cases $k = 0$, $k - \beta = 0$. There are also theorems corresponding to Theorems 40–44 in which "Lip" is replaced by "Lip*" and "O" by "o". These we may leave to the reader.

6. Further theorems concerning the means M_p.

6.1. Theorem 45. *Suppose that* $p > 0$, $a \geqq 0$, *and*

$$M_p(f) \leqq C(1-r)^{-a}.$$

If $\beta > -a - \frac{1}{p}$, *then*

$$|f^{\beta}(z)| \leqq KC(1-r)^{-a-\beta-\frac{1}{p}}.$$

If $-a-1-\frac{1}{p} < \beta < -a-\frac{1}{p}$, *then* $f^{\beta}(z)$ *belongs to*

$$\mathrm{Lip}\left(-\alpha-\frac{1}{p}-\beta\right).$$

If $a=0$, *we may replace* O *by* o *and* Lip *by* Lip* *in the conclusions.*

This theorem follows by combination of Theorem 27 with those of the preceding section.

The special case

$$a=0, \quad \beta=-\alpha<0, \quad p>1, \quad \frac{1}{p}<\alpha<1+\frac{1}{p},$$

for example, gives a theorem substantially equivalent to Theorem 12 of 7.

6.2. Theorem 46. *If*

$$p>0, \quad a\geqq 0, \quad a+\beta>0,$$

and

$$M_p(f) \leqq C(1-r)^{-a}, \tag{6.21}$$

then

$$M_p(f^{\beta}) \leqq KC(1-r)^{-a-\beta}, \tag{6.22}$$

where $K=K(p,a,\beta)$. *If* $a=0$, $\beta>0$, *then also*

$$M_p(f^{\beta}) = o\left((1-r)^{-\beta}\right). \tag{6.23}$$

(i) Suppose first that $p \geqq 1$. We distinguish the sub-cases

$$p\geqq 1, \quad \beta>0; \tag{ia}$$

$$p\geqq 1, \quad \beta=-\alpha, \quad 0<\alpha<a. \tag{ib}$$

When $\beta=0$ there is naturally nothing to prove.

(ia) In this case we have

$$f^{\beta}(z) = \frac{\Gamma(1+\beta)}{2\pi i}\int_{\Gamma_z}\frac{f(u)\,du}{(u-z)^{1+\beta}} \tag{6.24}$$

$$= \frac{\Gamma(1+\beta)}{2\pi i}\left(\int_{\gamma_z}+\int_{\Gamma_{1,z}}+\int_{\Gamma_{2,z}}\right) = F_1+F_2+F_3.$$

We take ϱ, the radius of Γ_z, to be $\frac{1}{2}(1+r)$. Then plainly

$$|F_2| \leqq K\int_0^{\varrho}|f(te^{i\theta+i\pi})|\,dt \leqq KC\int_0^{\varrho}(1-t)^{-a-\frac{1}{p}}\,dt,$$

by Theorem 27, (2.13); and F_2 is bounded if $\alpha+\frac{1}{p}<1$, of order

$$(1-\varrho)^{1-\alpha-\frac{1}{p}}=K(1-r)^{1-\alpha-\frac{1}{p}}$$

in the contrary case. It is plain in either case that

$$|F_2|\leqq KC(1-r)^{-\alpha}, \quad M_p(F_2)\leqq KC(1-r)^{-\alpha},$$

and *a fortiori*

$$M_p(F_2)\leqq KC(1-r)^{-\alpha-\beta}. \tag{6.25}$$

Similarly for F_3.

Next

$$\int_{-\pi}^{\pi}|F_1|^p\,d\theta\leqq K\int_{-\pi}^{\pi}d\theta\left(\int_{-\pi}^{\pi}\frac{|f(\varrho e^{i\varphi+i\theta})|\,d\varphi}{|\varrho e^{i\varphi}-r|^{1+\beta}}\right)^p. \tag{6.26}$$

If $p>1$, we can choose β_1 and β_2 so that

$$\beta_1+\beta_2=1+\beta, \quad p\beta_1>1, \quad p'\beta_2>1.$$

Then, by Hölder's inequality,

$$\left(\int_{-\pi}^{\pi}\frac{|f(\varrho e^{i\varphi+i\theta})|\,d\varphi}{|\varrho e^{i\varphi}-r|^{1+\beta}}\right)^p\leqq\int_{-\pi}^{\pi}\frac{|f(\varrho e^{i\varphi+i\theta})|^p\,d\varphi}{|\varrho e^{i\varphi}-r|^{p\beta_1}}\left(\int_{-\pi}^{\pi}\frac{d\varphi}{|\varrho e^{i\varphi}-r|^{p'\beta_2}}\right)^{p-1};$$

and the second factor is not greater than

$$K(\varrho-r)^{(1-p'\beta_2)(p-1)}\leqq K(1-r)^{p-1-p\beta_2}.$$

Also

$$\int_{-\pi}^{\pi}d\theta\int_{-\pi}^{\pi}\frac{|f(\varrho e^{i\varphi+i\theta})|^p\,d\varphi}{|\varrho e^{i\varphi}-r|^{p\beta_1}}=\int_{-\pi}^{\pi}\frac{d\varphi}{|\varrho e^{i\varphi}-r|^{p\beta_1}}\int_{-\pi}^{\pi}|f(\varrho e^{i\varphi+i\theta})|^p\,d\theta$$

$$\leqq KC^p(1-\varrho)^{-p\alpha}\int_{-\pi}^{\pi}\frac{d\varphi}{|\varrho e^{i\varphi}-r|^{p\beta_1}}$$

$$\leqq KC^p(1-\varrho)^{-p\alpha}(\varrho-r)^{1-p\beta_1}\leqq KC^p(1-r)^{1-p\alpha-p\beta_1}.$$

Hence

$$\int_{-\pi}^{\pi}|F_1|^p d\theta\leqq KC^p(1-r)^{p-1-p\beta_2}(1-r)^{1-p\alpha-p\beta_1}\leqq KC^p(1-r)^{-p\alpha-p\beta},$$

$$M_p(F_1)\leqq KC(1-r)^{-\alpha-\beta}; \tag{6.27}$$

and (6.22) follows from (6.24), (6.25) and (6.27).

The proof is simpler when $p=1$, no use of Hölder's inequality being required.

We have still to prove (6.23) when $a = 0$. In this case $f(z)$ has, by Lemma A(1), a boundary function $f(e^{i\theta})$ of L^p, and we may take $\varrho = 1$ in (6.24). The contributions of F_2 and F_3 are $O(1)$ and *a fortiori* $o((1-r)^{-\beta})$.

We choose δ so that

$$\frac{1}{2\pi}\int_{-\delta}^{\delta}|f(e^{i\varphi+i\theta})|^p d\varphi < \varepsilon$$

for all θ, and write (6.26) in the form

$$\int_{-\pi}^{\pi}|F_1|^p d\theta \leqq K\int_{-\pi}^{\pi}(J_1+J_2)^p d\theta \leqq K\left(\int_{-\pi}^{\pi}J_1^p d\theta + \int_{-\pi}^{\pi}J_2^p d\theta\right),$$

where

$$J_1 = \int_{-\delta}^{\delta}\frac{|f(e^{i\varphi+i\theta})|}{|e^{i\varphi}-r|^{1+\beta}}d\varphi, \qquad J_2 = \int_{|\varphi|>\delta}\frac{|f(e^{i\varphi+i\theta})|}{|e^{i\varphi}-r|^{1+\beta}}d\varphi.$$

Plainly J_2 is bounded, and our former argument, applied to J_1, will give

$$\int J_1^p d\theta \leqq K\varepsilon C^p(1-r)^{-p\beta}.$$

We thus obtain (6.23).

6.3. (i*b*) If $p \geqq 1$, $\beta = -\alpha$, $0 < \alpha < a$, we have

$$|f_a(re^{i\theta})| \leqq K\int_0^r (r-\varrho)^{\alpha-1}|f(\varrho e^{i\theta}|\,d\varrho,$$

$$\left(\int_{-\pi}^{\pi}|f_a(re^{i\theta})|^p d\theta\right)^{\frac{1}{p}} \leqq K\int_0^r (r-\varrho)^{\alpha-1}d\varrho\left(\int_{-\pi}^{\pi}|f(\varrho e^{i\theta})|^p d\theta\right)^{\frac{1}{p}}$$

$$\leqq K\int_0^r (r-\varrho)^{\alpha-1}(1-\varrho)^{-a}d\varrho$$

$$= K(1-r)^{\alpha-a}\int_1^{\frac{1}{1-r}}(x-1)^{\alpha-1}x^{-a}dx < K(1-r)^{\alpha-a}\int_1^{\infty} = K(1-r)^{\alpha-a},$$

which proves the theorem.

It may be observed that in this argument we do not use the regularity of $f(z)$; the argument is valid if, for example, $f(z)$ is an analytic function with branch points and $|f(z)|$ is one-valued.

It is also to be observed that we have really proved more than is asserted by (6.22), and in fact that

$$M_p(f_\alpha^*) = O((1-r)^{\alpha-a}). \tag{6.31}$$

We shall make use of these remarks later.

6.4. (ii) We now suppose $p<1$. It is plainly sufficient to prove the results (ii$a$) when $\beta=1$ and (ii$b$) when $\beta=-\alpha$ and $0<\alpha<a$.[18])

(ii$a$) Suppose that $p<1$, $\beta=1$, and, in the first instance, that $f(z)$ has no zeros in the unit circle. If we choose k so that $pk>1$, and write $f=g^k$, then

$$M_{pk}(g) \leqq C^{\frac{1}{k}}(1-r)^{-\frac{a}{k}};$$

and so, from (ia) above,

$$M_{pk}(g') \leqq K C^{\frac{1}{k}}(1-r)^{-\frac{a}{k}-1}.$$

But

$$\begin{aligned}\int |f'|^p d\theta &= K\int |g^{k-1}g'|^p d\theta \\ &\leqq K\left(\int |g|^{pk}d\theta\right)^{\frac{1}{k'}}\left(\int |g'|^{pk}d\theta\right)^{\frac{1}{k}} \\ &\leqq K C^p\left((1-r)^{-\frac{a}{k}}\right)^{\frac{pk}{k'}}\left((1-r)^{-\frac{a}{k}-1}\right)^{\frac{pk}{k}} = K C^p (1-r)^{-ap-p},\end{aligned}$$

which is (6.22). Plainly (6.23) may be proved similarly.

In the general case we may argue as in § 2.1 (ii). Taking r and ϱ as there, determining f_1 and f_2 so as to satisfy (2.16), and applying what we have proved to $f_1(z)=F_1(\varrho z)$, we obtain

$$M_p(f_1') \leqq 2KC_\varrho\left(1-\frac{r}{\varrho}\right)^{-1} \leqq KC(1-r)^{-a-1};$$

and the proof may be completed as before.

6.5. (ii$b$) There remains the case $p<1$, $\beta=-\alpha$, $0<\alpha<a$. We give two proofs, by arguments analogous to those of §§ 4.63 and 4.64.

---

[18]) We might have used this remark in case (i), and simplified the analysis of § 6.2 by supposing $\beta=1$. There, however, there is no great gain in simplicity, whereas here the simplification is forced upon us, since the argument used under ii$a$) depends essentially on the fact that β is an integer.

(i) Using the same notation as in § 4.63, and the transformation used already in § 5.1, we have

$$M_p^p(f_\alpha^*) \leqq K \sum_1^\infty 2^{-np\alpha} r^{p\alpha} \frac{1}{2\pi} \int_{-\pi}^{\pi} \mu_n^p \, d\theta$$

$$\leqq K \sum_1^\infty 2^{-np\alpha} r^{p\alpha} M_p^p(r_n, f)$$

$$\leqq K \int_0^r (r-\varrho)^{p\alpha-1} M_p^p(\varrho, f) \, d\varrho$$

$$\leqq KC^p \int_0^r (r-\varrho)^{p\alpha-1} (1-\varrho)^{-p\alpha} d\varrho$$

$$= KC^p (1-r)^{p\alpha-p\alpha} \int_1^{\frac{1}{1-r}} (x-1)^{p\alpha-1} x^{-p\alpha} dx < KC^p (1-r)^{p\alpha-p\alpha} \int_1^\infty$$

$$= KC^p (1-r)^{p\alpha-p\alpha},$$

which proves the theorem.

(ii) Suppose first that $\alpha < ap$. We can then choose s so that

$$\frac{1}{p} < s < \frac{a}{\alpha},$$

and we write

$$f = g^s.$$

In general g is not regular in the unit circle, having branch points at the zeros of f; but $|g(z)|$ is one-valued, and we can therefore apply (6.31) to g.

We have now

$$M_{sp}(g) \leqq C^{\frac{1}{s}} (1-r)^{-\frac{a}{s}},$$

and so, from (6.31),

$$M_{sp}(g_\alpha^*) \leqq KC^{\frac{1}{s}} (1-r)^{\alpha-\frac{a}{s}}. \tag{6.51}$$

Now

$$f_\alpha^*(re^{i\theta}) = K \int_0^r |g(\varrho e^{i\theta})|^s (r-\varrho)^{\alpha-1} d\varrho$$

$$\leqq K \mathfrak{G}^{s-1}(r,\theta) \int_0^r |g(\varrho e^{i\theta})| (r-\varrho)^{\alpha-1} d\varrho$$

$$= K \mathfrak{G}^{s-1}(r,\theta) \, g_\alpha^*(r,\theta),$$

where

$$\mathfrak{G}(r,\theta) = \operatorname*{Max}_{\varrho \leqq r} |g(\varrho e^{i\theta})|;$$

so that

$$\int f_\alpha^{*p} d\theta \leqq K \int \mathfrak{G}^{(s-1)p} g_\alpha^{*p} d\theta$$

$$\leqq K \Big(\int \mathfrak{G}^{sp} d\theta \Big)^{\frac{1}{s'}} \Big(\int g_\alpha^{*sp} d\theta \Big)^{\frac{1}{s}}.$$

But

$$\int \mathfrak{G}^{sp} d\theta = \int \mathfrak{F}^p d\theta \leqq K \int |f|^p d\theta \leqq K C^p (1-r)^{-p\alpha},$$

by Theorem 32; and

$$\int g_\alpha^{*sp} d\theta \leqq K C^p (1-r)^{sp\alpha - p\alpha},$$

by (6.51). Hence

$$\int f_\alpha^{*p} d\theta \leqq K C^p (1-r)^{-\frac{p\alpha}{s'}} (1-r)^{p\alpha - \frac{p\alpha}{s}} = K(1-r)^{p\alpha - p\alpha},$$

which proves (6.22).

We have thus proved that (6.21) implies

$$M_p(f_{\alpha_1}) \leqq K C (1-r)^{\alpha_1 - \alpha}$$

for $0 < \alpha_1 < p\alpha$. A repetition of the argument shows that

$$M_p(f_{\alpha_2}) \leqq K C (1-r)^{\alpha_2 - \alpha}$$

for $0 < \alpha_2 - \alpha_1 < p(\alpha - \alpha_1)$, so that (6.22) holds for $0 < \alpha < p\alpha + (1-p)p\alpha$. Another repetition shows that it holds for $\alpha < p\alpha + (1-p)p\alpha + (1-p)^2 p\alpha$; and some finite number of repetitions that it holds for any α less than

$$(1 + (1-p) + (1-p)^2 + \ldots) p\alpha = \alpha.$$

This completes the proof of Theorem 46.

6.6. We can now obtain a very general theorem by combination of the theorems proved already.

Theorem 47. *If*

(6.61) $$0 < p \leqq q, \quad \alpha > 0, \quad \beta > -\alpha - \frac{1}{p} + \frac{1}{q},$$

or

(6.62) $$0 < p \leqq q, \quad \alpha = 0, \quad \beta \geqq -\frac{1}{p} + \frac{1}{q},$$

and

(6.63) $$M_p(f) \leqq C(1-r)^{-\alpha},$$

then

(6.64) $$M_q(f^\beta) \leqq K C (1-r)^{-\alpha - \beta - \frac{1}{p} + \frac{1}{q}},$$

where $K = K(p, q, a, \beta)$. *The result is not necessarily true when* $a > 0$, $\beta = -a - \frac{1}{p} + \frac{1}{q}$. *But when* $a = 0$, $\beta > -\frac{1}{p} + \frac{1}{q}$, *we have also*

$$M_q(f^\beta) = o\left((1-r)^{-\beta-\frac{1}{p}+\frac{1}{q}}\right). \tag{6.65}$$

(i) Suppose first that $a > 0$, $\beta = -\alpha$, so that

$$\alpha < a + \frac{1}{p} - \frac{1}{q}.$$

If

$$\gamma = \frac{1}{p} - \frac{1}{q}, \quad q = \frac{p}{1-p\gamma}$$

we have

$$M_q(f_\gamma) \leqq KC(1-r)^{-a}$$

by Theorem 38, and therefore

$$M_q(f_\alpha) \leqq KC(1-r)^{\alpha-\gamma-a}$$

by Theorem 46 (6.22). This is equivalent to (6.64).

To prove the falsity of the result when

$$\alpha = a + \frac{1}{p} - \frac{1}{q}$$

take

$$f(z) = (1-z)^{-a-\frac{1}{p}}.$$

Then (6.63) is satisfied, but $f_\alpha(z)$ behaves like a constant multiple of $(1-z)^{-\frac{1}{q}}$, and $M_q(f_\alpha) \to \infty$.

(ii) Next suppose that $a = 0$. If

$$\alpha = \frac{1}{p} - \frac{1}{q}$$

the theorem reduces to Theorem 33. If

$$\alpha < \frac{1}{p} - \frac{1}{q}$$

the results follow from Theorem 33 and Theorem 46 (6.23).

7. Integrated Lipschitz conditions.

7.1. In 7 we defined the class $\mathrm{Lip}(k, p)$, where $p \geqq 1$, $0 < k \leqq 1$, as the class of functions $F(\theta)$ for which

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |F(\theta+h) - F(\theta-h)|^p d\theta \leqq C^p h^{pk},$$

where $h > 0$ and C is independent of h; and the class $\mathrm{Lip}^*(k, p)$, where $p \geqq 1$, $0 < k < 1$, as the class for which

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |F(\theta+h) - F(\theta-h)|^p d\theta = o(h^{pk}),$$

when $h \to 0$. It proved indifferent which of the differences

$$F(\theta+h) - F(\theta-h), \qquad F(\theta+h) - F(\theta), \qquad F(\theta) - F(\theta-h)$$

we selected.

We considered only the case $p \geqq 1$, and we do not propose to go beyond this case here, so that we have nothing essentially new to prove. There are however certain theorems which are required for an application which we propose to make of the results of this paper to Parseval's Theorem.

We may say that $f(z)$ belongs to $\mathrm{Lip}(k, p)$ if it belongs to L^p and its boundary function $f(e^{i\theta})$ belongs to $\mathrm{Lip}(k, p)$; or, what is the same thing, if

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(re^{i\theta+ih}) - f(re^{i\theta-ih})|^p d\theta \leqq C^p h^{pk},$$

where C is independent of h and r.

Theorem 48. *A necessary and sufficient condition that $f(z)$ should belong to $\mathrm{Lip}(k, p)$ is that*

$$M_p(f') = O((1-r)^{k-1}). \tag{7.11}$$

This is Theorem 3 of 8 [19]).

---

[19]) There is one point in the proof of this theorem where we were perhaps not sufficiently explicit. We took for granted on p. 626 the existence of the limit function $f(e^{i\theta})$ of an $f(z)$ satisfying (7.11). If (7.11) is satisfied we have (using Minkowski's inequality as in § 6.3, and assuming as we may that $f(0) = 0$),

$$\begin{aligned} M_p(f) &\leqq K\left(\int d\theta\left(\int_0^r |f'(\varrho e^{i\theta})|\, d\varrho\right)^p\right)^{\frac{1}{p}} \\ &\leqq K\int_0^r d\varrho\left(\int |f'(\varrho e^{i\theta})|^p d\theta\right)^{\frac{1}{p}} \leqq K\int_0^r M_p(f')\, d\varrho \\ &= O\left(\int_0^r \frac{d\varrho}{(1-\varrho)^{1-k}}\right) = O(1), \end{aligned}$$

so that f has a limit function of L^p.

7.2. Theorem 49. *If* $f(z)$ *belongs to* $\mathrm{Lip}(k, p)$, *and*

$$-1+k<\beta<k,$$

then $f^{\beta}(z)$ *belongs to* $\mathrm{Lip}(k-\beta, p)$.

This theorem, which is substantially equivalent to Theorems 25 and 26 of 7, follows at once from Theorems 46 and 48. We have in this way an alternative proof of the theorems referred to.

We conclude by stating without proof the theorem for the class $\mathrm{Lip}(k, p)$, which corresponds to Theorem 41.

Theorem 50. *If* $f(z)$ *belongs to* $\mathrm{Lip}(k, p)$, *then* $f(re^{i\theta})$, *regarded as a function of* r *on the radius* $0 \leqq r \leqq 1$, *belongs to* $\mathrm{Lip}(k, p)$, *and uniformly in* θ; *that is to say*

$$\int_h^1 |f(re^{i\theta})-f((r-h)e^{i\theta})|^p\,dr = O(h^{pk}),$$

uniformly in θ.

The converse inference, which would correspond to Theorem 50 as Theorem 42 corresponds to Theorem 41, is false. The function

$$f(z)=\sum 2^{-kn} z^{2^n} \qquad \left(0<k<k+\frac{1}{p}<1\right)$$

belongs to $\mathrm{Lip}\left(k+\frac{1}{p}, p\right)$ on a radius, but on the circle only to $\mathrm{Lip}(k, p)$.

References.

1. Hadamard, J.: Essai sur l'étude des fonctions données par leur développement de Taylor, Journal de Math. (4), 8 (1892), p. 101–186.
2. Hardy, G. H.: The mean value of the modulus of an analytic function, Proc. London Math. Soc. (2), 14 (1915), p. 269–277.
3. Hardy, G. H., and Littlewood, J. E: Some properties of fractional integrals, Proc. London Math. Soc. (2), 24 (1925), p. xxxii–xli (Records for 12 March 1925).
4. — — Some new properties of Fourier constants, Math. Ann. 97 (1926), p. 159–209.
5. — — Notes on the theory of series (V); On Parseval's Theorem, Proc. London Math. Soc. (2), 26 (1926), p. 287–294.
6. — — Notes on the theory of series (VI); Two inequalities, Journal London Math. Soc. 2 (1927), p. 196–201.
7. — — Some properties of fractional integrals (I), Math. Zeitschr. 27 (1928), p. 565–606.
8. — — A convergence criterion for Fourier series, Math. Zeitschr. 28 (1928), p. 612–634.
9. — — A maximal theorem with function-theoretic applications, Acta Math. 54 (1930), p. 81–116.
10. Landau, E.: Darstellung und Begründung einiger neuerer Ergebnisse der Funktionentheorie, 2. Aufl., Berlin 1929.
11. Littlewood, J. E.: On the mean values of power series, Proc. London Math. Soc. (2), 25 (1926), p. 328–337.

12. Pólya, G., und Szegö, G.: Aufgaben aus der Analysis (I), Berlin 1925.
13. Riesz, F.: Über die Randwerte einer analytischen Funktion, Math. Zeitschr. 18 (1923), p. 87–95.
14. — Sur les fonctions subharmoniques et leur rapport à la théorie du potentiel, Acta Math. 48 (1926), p. 329–343.
15. — F. und M.: Über die Randwerte einer analytischen Funktion, Compte rendu du IV Congrès des mathématiciens Scandinaves (1920), p. 27–44.
16. — M.: Sur les fonctions conjuguées, Math. Zeitschr. 27 (1927), p. 218–244.
17. Weyl, H.: Bemerkungen zum Begriff des Differentialquotienten gebrochener Ordnung, Vierteljahrsschrift d. naturf. Ges. in Zürich 62 (1917), p. 296–302.

(Eingegangen am 27. April 1931.)

Addendum.

We take this opportunity of correcting a mistake in 7 pointed out to us by Miss M. L. Cartwright. The proof of Theorem 24 in the case $p = 1$ (p. 604–605) is incomplete until the restriction on the set e is removed.

The set E in which φ_n does not tend to f is of measure zero. We define f^* as follows: in the complementary set CE, $f^* = f$; in E

$$f^*(x) = \overline{\lim} f(z)$$

when $z \to x$ through points of CE. Then f^* is equivalent to f, and therefore belongs to $\mathrm{Lip}(1, 1)$. The argument of the text, applied to f^*, shows that

$$\sum_e |f^*(\eta) - f^*(\xi)| \leqq C$$

for any set e whose extremities lie in CE. But if x lies in E, $f^*(x)$ is a limit of $f^*(z)$ for values of z lying in CE and tending to x, so that a continuity argument is sufficient to remove the restriction on e.

Ein Konvergenzsatz für mehrvariablige Fouriersche Integrale.

Von

S. Bochner in München.

§ 1.

Einer nach Lebesgue integrierbaren Funktion $f(x, y)$, welche in jeder Variablen die Periode 2π hat, kann man eine Fouriersche Reihe zuordnen,

$$f(x, y) \sim \sum\sum c_{\mu\nu} e^{i(\mu x + \nu y)}. \tag{1}$$

Die Koeffizienten dieser Reihe sind definiert durch

$$c_{\mu\nu} = \frac{1}{(2\pi)^2} \int_0^{2\pi}\int_0^{2\pi} f(\xi, \eta)\, e^{-i(\mu\xi + \nu\eta)}\, d\xi\, d\eta, \tag{2}$$

und ihre Partialsummen

$$S_{mn}(x, y) = \sum_{-m}^{m}\sum_{-n}^{n} c_{\mu\nu} e^{i(\mu x + \nu y)} \tag{3}$$

haben den Wert

$$S_{mn}(x, y) = \frac{1}{\pi^2} \int_0^{\pi}\int_0^{\pi} f(x + 2u, y + 2v)\, \frac{\sin(2m+1)u}{\sin u} \cdot \frac{\sin(2n+1)v}{\sin v}\, du\, dv. \tag{4}$$

Wenn man nach Bedingungen fragt, unter denen in einem Punkte x_0, y_0 die Doppelfolge S_{mn} konvergent ist (etwa gegen den Wert $f(x_0, y_0)$), so stößt man bekanntlich auf die Schwierigkeit, daß im Gegensatz zum einvariabligen Falle die Konvergenz der Partialsummen nicht durch das bloße Verhalten von $f(x_0 + 2u, y_0 + 2v)$ in einer beliebigen Umgebung $u^2 + v^2 \leqq \varepsilon^2$ determiniert wird. Vielmehr hängt der Wert von $S_{mn}(x_0, y_0)$ für große m und n wesentlich davon ab, wie sich $f(x_0 + 2u, y_0 + 2v)$ in den beiden Streifen $|u| \leqq \varepsilon$, $0 \leqq v < \pi$ und $|v| \leqq \varepsilon$, $0 \leqq u < \pi$ verhält. Daher kommt es, daß die relativ scharfen Konvergenzkriterien sich im Wortlaut etwas umständlich gestalten[1]).

[1]) Vgl. L. Tonelli, Serie trigonometriche, (1928), Kap. IX, außerdem H. Geiringer, Monatshefte f. Math. u. Physik 29 (1918), S. 65—114.

Dieselbe Komplizierung, sogar etwas stärker, stellt sich ein, wenn man nicht Fouriersche *Reihen*, sondern Fouriersche *Integrale* betrachtet. Eine Funktion $f(x, y)$ sei in der ganzen Ebene definiert und nach Lebesgue integrierbar und besitze ein endliches Integral

$$\int_{-\infty}^{\infty}\int_{-\infty}^{+\infty} |f(\xi, \eta)|\, d\xi\, d\eta\,.^{2}) \tag{5}$$

Den Koeffizienten (2) entspricht die Transformierte

$$\Gamma(\alpha, \beta) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\xi, \eta)\, e^{-i(\alpha\xi+\beta\eta)}\, d\xi\, d\eta\,, \tag{6}$$

der Doppelreihe (1) das Doppelintegral

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Gamma(\alpha, \beta)\, e^{i(\alpha x+\beta y)}\, d\alpha\, d\beta\,, \tag{7}$$

und der Partialsumme (3) das Partialintegral

$$S_{ab}(x, y) = \int_{-a}^{a}\int_{-b}^{b} \Gamma(\alpha, \beta)\, e^{i(\alpha x+\beta y)}\, d\alpha\, d\beta\,, \tag{8}$$

welches den Wert hat

$$S_{ab} = \frac{1}{\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x+u, y+v)\, \frac{\sin au}{u}\cdot\frac{\sin bv}{v}\, du\, dv\,. \tag{9}$$

Damit $S_{ab}(x_0, y_0)$ für große a und b konvergieren soll[3]), bedarf es Einschränkungen der Funktion $f(x_0+u, y_0+v)$ in den *unendlichen* Streifen: $|u| \leqq \varepsilon$, $-\infty < v < +\infty$ und $|v| \leqq \varepsilon$, $-\infty < u < +\infty$. Die Analoga zu den vorliegenden relativ scharfen Kriterien für die Reihe (1) würden sich noch umständlicher gestalten als diese selbst (und für mehr- als zweivariablige Funktionen ziemlich undurchsichtig werden).

Demgegenüber dürfte es von theoretischem (wenn auch kaum von praktischem) Interesse sein, daß man ein recht prägnantes Konvergenzkriterium angeben kann, falls man nicht die „rechteckigen" Partialintegrale (8), sondern die „kreisförmigen" Partialintegrale

$$S_R(x, y) = \int\int_{\alpha^2+\beta^2 \leqq R^2} \Gamma(\alpha, \beta)\, e^{i(\alpha x+\beta y)}\, d\alpha\, d\beta \tag{10}$$

zugrunde legt. Dann besteht der folgende

---

[2]) Ohne diese Endlichkeitsvoraussetzung würden die Definition und die weitere Handhabung der Transformierten $\Gamma(\alpha, \beta)$ auf erhebliche Schwierigkeiten stoßen.

[3]) Und zwar gegen $f(x_0, y_0)$ oder, falls $f(xy)$ an der Stelle x_0, y_0 unstetig ist, gegen einen passend definierten Grenzwert.

Konvergenzsatz. *Damit der Limes*

$$\lim_{R\to\infty} S_R(x_0, y_0)$$

vorhanden und endlich ist, ist hinreichend, daß der kreisförmige Mittelwert von $f(x\,y)$ um den Punkt x_0, y_0 herum, nämlich die Funktion

$$\bar{f}(t) = \frac{1}{2\pi}\int_0^{2\pi} f(x_0 + t\cos\varphi,\ y_0 + t\sin\varphi)\,d\varphi, \tag{11}$$

als Funktion der Distanz t betrachtet im Gesamtintervall $0 < t < +\infty$ von beschränkter Variation ist. Es ist dann

$$\lim S_R(x_0, y_0) = \bar{f}(0) = \lim_{t\to 0}\bar{f}(t).$$

Über die Existenz der Funktion $\bar{f}(t)$ ist folgendes zu bemerken. Die Funktion $f(x, y)$ ist nach Voraussetzung überall im Endlichen integrierbar. Führt man die Polarkoordinaten $x = x_0 + t\cos\varphi$, $y = y_0 + t\sin\varphi$ ein[4]), so ergibt sich nach dem Satz von Fubini, daß das Integral (11) für fast alle positiven Werte von t vorhanden ist. Und unsere Konvergenzbedingung lautet dahin, daß die so resultierende Funktion $\bar{f}(t)$ nach passender Abänderung in einer Nullmenge eine beschränkte Gesamtvariation haben soll.

In § 2 werden wir diesen Satz beweisen, und in § 3 ein Analogon für n-variablige Funktionen aufstellen.

§ 2.

Wenn man in (10) für $\Gamma(\alpha, \beta)$ den Wert (6) einsetzt, und auf Grund der Konvergenz von (5) die Reihenfolge der Integrationen vertauscht, so erhält man

$$S_R(x_0, y_0) = \frac{1}{(2\pi)^2}\int_{-\infty}^{+\infty}\int f(x_0 + \xi, y_0 + \eta)\,H_R(\xi, \eta)\,d\xi\,d\eta, \tag{12}$$

wobei

$$H_R(\xi, \eta) = \int_{\alpha^2+\beta^2\leqq R^2}\int e^{-i(\xi\alpha+\eta\beta)}\,d\alpha\,d\beta.$$

Jetzt führen wir Polarkoordinaten ein,

$$\xi = t\cos\varphi, \qquad \eta = t\sin\varphi,$$

$$\alpha = \gamma\cos\psi, \qquad \beta = \gamma\sin\psi,$$

und erhalten

$$H_R(\xi, \eta) = \int_0^R\int_0^{2\pi} e^{-it\gamma\cos(\psi-\varphi)}\,\gamma\,d\gamma\,d\psi.$$

---

[4]) Wegen der Zulässigkeit dieser Variablentransformation vgl. etwa E. W. Hobson, Functions of a Real Variable 1 (1921), S. 595—599.

Nun ist

$$\int_0^{2\pi} e^{-iz\cos(\psi-\varphi)}\,d\psi = 2\pi J_0(z),$$

wobei wir, wie üblich, mit $J_\mu(z)$ die Besselsche Funktion μ-ter Ordnung bezeichnen. Daher ist

$$\begin{aligned} H_R(\xi,\eta) &= 2\pi\int_0^R J_0(t\gamma)\,\gamma\,d\gamma \\ &= \frac{2\pi}{t^2}\int_0^{Rt} J_0(\gamma)\,\gamma\,d\gamma \\ &= \frac{2\pi}{t^2}\int_0^{Rt}\frac{d}{d\gamma}[J_1(\gamma)\gamma]\,d\gamma \\ &= \frac{2\pi R}{t}J_1(Rt). \end{aligned}$$

Wenn man dies in das Doppelintegral (12) einsetzt und nunmehr nach dem Satz von Fubini zuerst nach φ und dann nach t integriert, so erhält man

$$R\int_0^\infty \bar{f}(t)\,J_1(Rt)\,dt,$$

und daher endgültig

$$S_R(x_0,y_0) = \int_0^\infty \bar{f}\left(\frac{t}{R}\right)J_1(t)\,dt.$$

Wegen

$$J_1(t) = \sqrt{\frac{2}{\pi}}\cdot\frac{1}{\sqrt{t}}\cos\left(t-\frac{3\pi}{4}\right)+O(t^{-3/2})$$

ist die Funktion

$$\varepsilon(N) = \underset{\varrho>0}{\text{Ob. Gr.}}\int_N^{N+\varrho} J_1(t)\,dt$$

überall endlich und für $N\to\infty$ nach Null konvergent. Nach dem zweiten Mittelwertsatz ist

$$\left|\int_N^\infty \bar{f}\left(\frac{t}{R}\right)J_1(t)\,dt\right| \leqq (|\bar{f}(\infty)| + V_{\bar{f}}(0,\infty))\cdot\varepsilon(N).$$

Andererseits ist

$$\int_0^\infty J_1(t)\,dt = 1.$$

Hieraus folgt in bekannter Weise, daß

$$S_R(x_0,y_0)\to\bar{f}(0).$$

§ 3.

Nunmehr legen wir den n-dimensionalen Kartesischen Raum zugrunde $n \geqq 2$. Mit K_R werden wir die n-dimensionale Kugel

$$\alpha_1^2 + \alpha_2^2 + \ldots + \alpha_n^2 \leqq R^2$$

bezeichnen, mit S_R die $(n-1)$-dimensionale Sphäre

$$\xi_1^2 + \xi_2^2 + \ldots + \xi_n^2 = R^2,$$

und mit ω bzw. $d\omega$ den Inhalt bzw. das Inhaltselement der Sphäre S_1. Weiterhin werden wir Größen $c_1, c_2, \ldots$ einführen, von denen jede ausschließlich von der Dimensionszahl n abhängt.

Eine Funktion

$$f(x) \equiv f(x_1, x_2, \ldots, x_n)$$

sei überall im Raume definiert und nach Lebesgue integrierbar und besitze ein endliches Integral

$$\int_{-\infty}^{+\infty} \ldots \int |f(x)|\, dx_1 \ldots dx_n.$$

Wir bilden die Transformierte

$$\Gamma(\alpha) = \frac{1}{(2\pi)^n} \int_{-\infty}^{+\infty} \ldots \int f(\xi)\, e^{-i(\alpha_1\xi_1 + \ldots + \alpha_n\xi_n)}\, d\xi_1 \ldots d\xi_n$$

und mit ihr die Partialintegrale

$$S_R(x) = \int_{K_R} \ldots \int \Gamma(\alpha)\, e^{i(x_1\alpha_1 + \ldots + x_n\alpha_n)}\, d\alpha_1 \ldots d\alpha_n.$$

n-dimensionaler Konvergenzsatz. *Damit die Partialintegrale*

(13) $$S_R(x_1^0, x_2^0, \ldots, x_n^0)$$

für $R \to \infty$ gegen eine endliche Zahl konvergieren, ist hinreichend, daß der sphärische Mittelwert von $f(x)$ um den Punkt $x_1^0, x_2^0, \ldots, x_n^0$ herum, nämlich die Funktion

$$\bar{f}(t) = \frac{1}{\omega} \int_{S_1} \ldots \int f(x^0 + t\xi)\, d\omega,$$

als Funktion der Distanz t betrachtet von der folgenden Beschaffenheit ist: Bezeichnet man mit ν die Größe $\frac{n-2}{2}$, falls n gerade ist, und die Größe $\frac{n-1}{2}$, falls n ungerade ist, so besitzen die Funktionen

$$\bar{f}(t) \quad \text{und} \quad t^\nu \bar{f}(t)$$

Ableitungen bis zur ν-ten Ordnung, sind mitsamt diesen Ableitungen beschränkt, und die Funktion

$$\frac{d^\nu}{dt^\nu}\{t^\nu \bar{f}(t)\}$$

ist ein Intervall $0 < t < +\infty$ von beschränkter Variation. Der Limes von (13) *hat dann den Wert*

$$\bar{f}(0) = \lim_{t \to 0} \bar{f}(t).$$

Über die Existenz von $\bar{f}(t)$ gilt eine analoge Bemerkung wie in § 1.

Beweis. Wir schreiben, analog wie in § 2,

$$S_R(x^0) = c_1 \int_{-\infty}^{+\infty} \dots \int f(x^0 + \xi)\, H_R(\xi)\, d\xi_1 \dots d\xi_n, \tag{14}$$

wobei

$$H_R(\xi) = \int_{K_R} \dots \int e^{-i(\xi_1 \alpha_1 + \dots + \xi_n \alpha_n)}\, d\alpha_1 \dots d\alpha_n.$$

Wählt man im Raume $\alpha_1 \dots \alpha_n$ um den Punkt $0, \dots, 0$ herum ein neues orthogonales Koordinatensystem, so daß die neue α_1-Achse im Raume der alten Koordinaten die Richtungskosinusse

$$\xi_1 : \xi_2 : \dots : \xi_n$$

hat, und führt man die Größe

$$t = \sqrt{\xi_1^2 + \dots + \xi_n^2}$$

ein, so erhält man

$$H_R(\xi) = \int_{K_R} \dots \int e^{-it\alpha_1}\, d\alpha_1 \dots d\alpha_n.$$

Jetzt führen wir Polarkoordinaten $\gamma, \psi, \psi_1, \dots, \psi_{n-2}$ ein, die so beschaffen sind, daß γ das Intervall $(0, \infty)$, ψ das Intervall $(0, 2\pi)$ durchläuft, und daß

$$\alpha_1 = \gamma \cos\psi$$

$$\pm\sqrt{\alpha_2^2 + \dots + \alpha_n^2} = \gamma \sin\psi.$$

Dann ist

$$\frac{\partial(\alpha_1, \alpha_2, \dots, \alpha_n)}{\partial(\gamma, \psi, \psi_1, \dots, \psi_{n-2})} = \gamma^{n-1} (\sin\psi)^{n-2}\, D(\psi_2, \dots, \psi_n),$$

wo D eine von γ und ψ unabhängige Funktion ist. Daher ist

$$H_R(\xi) = c_2 \int_0^R \int_0^{2\pi} e^{-it\gamma\cos\psi} \gamma^{n-1} |\sin\psi|^{n-2}\, d\gamma\, d\psi.$$

Nun ist

$$\int_0^{2\pi} e^{-iz\cos\psi} |\sin\psi|^{n-2}\, d\psi = c_3\, z^{-\mu} J_\mu(z),$$

wobei

$$\mu = \frac{n-2}{2}.$$

Also ist

$$\begin{aligned} H_R(\xi) &= c_4 \int_0^R (t\gamma)^{-\mu} J_\mu(t\gamma)\, \gamma^{n-1}\, d\gamma \\ &= c_4\, t^{-n} \int_0^{Rt} J_\mu(\gamma)\, \gamma^{\mu+1}\, d\gamma \\ &= c_4\, t^{-n} \int_0^{Rt} \frac{d}{d\gamma}\left(J_{\mu+1}(\gamma)\, \gamma^{\mu+1}\right) d\gamma \\ &= c_4 \left(\frac{R}{t}\right)^{\mu+1} J_{\mu+1}(tR). \end{aligned}$$

Wenn man dies in das Integral (14) einsetzt und Polarkoordinaten einführt, so erhält man

$$c_5 \int_0^\infty \bar{f}(t)\, R^{\mu+1} t^\mu J_{\mu+1}(Rt)\, dt$$

und daher endgültig

$$S_R(x^0) = c_5 \int_0^\infty \bar{f}\left(\frac{t}{R}\right) t^\mu J_{\mu+1}(t)\, dt.$$

Die Funktion $J_{\mu+1}(t)$ hat für jedes μ eine asymptotische Entwicklung nach den Funktionen

$$\frac{e^{\pm it}}{\sqrt{t}}\, t^{-k} \qquad (k = 0, 1, 2, \ldots).$$

Daraus ergibt sich unschwer, daß auch jede der Funktionen

$$H_1(t) = \int_t^\infty J_{\mu+1}(t)\, dt,$$

$$H_{i+1}(t) = \int_t^\infty H_i(t)\, dt \qquad (i = 1, 2, 3, \ldots),$$

eine derartige Entwicklung besitzt. Hieraus folgt, daß $H_i(t)$ für $t \to \infty$ gegen Null konvergiert und eine Abschätzung der Form

$$\left| \int_N^{N+\varrho} H_i(t)\, dt \right| \leqq \varepsilon_i(N) \qquad (\varrho > 0)$$

zuläßt, wo

$$\lim_{N \to \infty} \varepsilon_i(N) = 0.$$

Nunmehr nehmen wir an, daß n eine grade, also μ eine ganze Zahl ist. Wir setzen

$$S_R(x^0) = c_5\, R^\mu \int_0^\infty g\left(\frac{t}{R}\right) J_{\mu+1}(t)\, dt,$$

wo

$$g(t) = t^{\mu} \bar{f}(t),$$

und hieraus erhält man nach (zulässiger) μ-maliger partieller Integration

$$S_R(x^0) = c_6 \int_0^{\infty} g^{(\mu)}\left(\frac{t}{R}\right) H_{\mu}(t)\, dt.$$

Eine analoge Überlegung wie in § 2 ergibt, daß

$$\lim_{R \to \infty} S_R(x^0)$$

vorhanden ist, und den Wert hat

$$c_7\, g^{(\mu)}(0).$$

Wegen der Beschränktheit der Ableitung von $\bar{f}(t)$ ist $g^{(\mu)}(0) = \bar{f}(0)$, daher hat unser Limes den Wert

$$c_7 \bar{f}(0).$$

Bleibt nur noch nachzuweisen, daß $c_7 = 1$. Das erhält man am einfachsten, wenn man die spezielle Funktion einsetzt

$$f(x_1, \ldots, x_n) = e^{-(x_1^2 + \ldots + x_n^2)}$$

und sie im Punkte $x_\nu^0 = 0$ betrachtet. In diesem Falle verifiziert man direkt, daß

$$f(0, \ldots, 0) = \int_{-\infty}^{+\infty} \ldots \int \Gamma(\alpha)\, d\alpha_1 \ldots d\alpha_n. \tag{15}$$

Da aber in diesem Falle das Integral (15) absolut konvergiert, ist auch

$$f(0, \ldots, 0) = \lim_{R \to \infty} S_R(0, \ldots, 0).$$

Wegen $f(0, \ldots, 0) = 1$ ist daher $c_7 = 1$.

Ist n ungerade, so setze man $\nu = \mu + \frac{1}{2}$ und

$$g(t) = t^{\nu} \bar{f}(t),$$

$$H_1(t) = \int_t^{\infty} \frac{J_{\mu+1}(t)}{\sqrt{t}}\, dt,$$

$$H_{i+1}(t) = \int_t^{\infty} H_i(t)\, dt \qquad (i = 1, 2, 3, \ldots).$$

Man schließt zu Ende wie oben, nur mit dem Unterschied, daß jetzt die Funktionen $H_i(t)$ eine asymptotische Entwicklung nach den Funktionen

$$\frac{e^{\pm it}}{t}\, t^{-k} \qquad (k = 0, 1, 2, \ldots)$$

haben.

(Eingegangen am 15. März 1931.)

Teilerprobleme.

Zweite Abhandlung.

Von

Arnold Walfisz in Radość (Polen).

Einleitung.

Es bezeichne $\sigma(n)$ die Summe der reziproken Teiler einer natürlichen Zahl n

$$\sigma(n) = \sum_{d|n} \frac{1}{d},$$

also $n\sigma(n)$ die Teilersumme

$$n\sigma(n) = \sum_{d|n} d.$$

Ich betrachte ferner die beiden summatorischen Funktionen

$$S_0(x) = \sum_{n \leqq x} \sigma(n), \qquad S(x) = \sum_{n \leqq x} n\sigma(n)$$

(falls die untere Summationsgrenze nicht explizit bezeichnet wird, ist sie stets Eins) und setze

$$S_0(x) = \frac{\pi^2}{6} x - \frac{1}{2} \log x + T_0(x), \tag{1}$$

$$S(x) = \frac{\pi^2}{12} x^2 + T(x). \tag{2}$$

Schließlich sei $A_4(x)$ die Anzahl der Gitterpunkte im Innern und auf der Oberfläche der vierdimensionalen Kugel $u_1^2 + u_2^2 + u_3^2 + u_4^2 \leqq x$ und

$$A_4(x) = \frac{\pi^2}{2} x^2 + \mathsf{P}_4(x). \tag{3}$$

In der ersten gleichnamigen Abhandlung[1]) habe ich mit Hilfsmitteln aus der Theorie Diophantischer Approximationen gezeigt, daß

$$T_0(x) = O\left(\frac{\log x}{\log\log x}\right),$$

$$T(x) = O\left(\frac{x \log x}{\log\log x}\right),$$

$$\mathsf{P}_4(x) = O\left(\frac{x \log x}{\log\log x}\right)$$

---

[1]) Mathematische Zeitschrift 26 (1927), S. 66–88. Im folgenden mit I zitiert.

ist. Schärfere Abschätzungen zu erzielen, dürfte sehr schwer sein; daher liegt es nahe, die betreffenden Fehlerquadratintegrale zu untersuchen.

Für diese soll im folgenden

$$\text{(I)} \qquad \int_0^x T_0^2(u)\,du = \left(\left(\frac{\gamma+\log 2\pi}{2}\right)^2 + \frac{5\pi^2}{144}\right)x + O\left(x^{\frac{1}{2}}\log x\right),$$

$$\text{(II)} \qquad \int_0^x T^2(u)\,du = \frac{36+5\pi^2}{432}x^3 + O\left(x^{\frac{5}{2}}\log x\right),$$

$$\text{(III)} \qquad \int_0^x \mathsf{P}_4^2(u)\,du = \frac{2}{3}\pi^2 x^3 + O\left(x^{\frac{5}{2}}\log x\right)$$

nachgewiesen werden. Hierbei ist mit γ die Eulersche Konstante gemeint. Daß

$$\text{(4)} \qquad \int_0^x \mathsf{P}_4^2(u)\,du = O(x^3)$$

ist, hat in viel allgemeinerem Rahmen kürzlich V. Jarník gezeigt[2]).

Die Beweise von (I), (II) und (III) erfordern nur einfache reell-analytische Überlegungen, sind aber rechnerisch etwas langwierig.

Folgende Festsetzungen mögen für die ganze vorliegende Arbeit gelten: Mit B bezeichne ich unterschiedslos reelle Zahlen, die von allem Möglichem abhängen dürfen, absolut genommen jedoch unterhalb numerischer Konstanten liegen. Es sei stets

$$R(u) = u - [u], \qquad \psi(u) = R(u) - \frac{1}{2}.$$

Die nicht kleinere der beiden Zahlen u und v heiße $\mathrm{Max}(u; v)$; dagegen bedeute (u, v) den größten gemeinsamen Teiler der beiden ganzen Zahlen u und v.

Summen wie

$$\sum_{a \leqq b \leqq \sqrt{x}}$$

laufen bald über a und b, bald nur über b. Jedesmal ergibt sich aber sofort aus dem Zusammenhange, was gemeint ist.

Die fetten Zahlen geben Nummern von Formeln an, die an den betreffenden Stellen benutzt werden, sind also als Hinweise zu verstehen.

§ 1.

Beweis von (I).

Hilfssatz 1:

$$\text{(5)} \qquad -T_0(x) = \frac{1}{2}(\gamma + \log 2\pi) + \sum_{a \leqq \sqrt{x}} \frac{1}{a}\psi\left(\frac{x}{a}\right) + \frac{1}{x}\sum_{b \leqq \sqrt{x}} b\,\psi\left(\frac{x}{b}\right) + B x^{-\frac{1}{2}} \qquad (x \geqq 1).$$

[2]) V. Jarník „Über die Mittelwertsätze der Gitterpunktlehre“ [Mathematische Zeitschrift **33** (1931), S. 62–84] gibt die zu (4) analoge Abschätzung für Ellipsoide $\alpha_1 u_1^2 + \alpha_2 u_2^2 + \alpha_3 u_3^2 + \alpha_4 u_4^2 \leqq x$ mit beliebigen positiven α.

Beweis: Für $1 \leqq x < 2$ lautet die Behauptung

$$\frac{\pi^2}{6}x - \frac{1}{2}\log x - 1 = \frac{1}{2}(\gamma + \log 2\pi) + \psi(x) + \frac{1}{x}\psi(x) + Bx^{-\frac{1}{2}} \quad (1).$$

Da diese Beziehung klar ist, soll jetzt $x \geqq 2$ angenommen werden.

$$S_0(x) = \sum_{n \leqq x} \sum_{a \mid n} \frac{1}{a} = \sum_{ab \leqq x} \frac{1}{a}$$

$$(6) \qquad = \sum_{\substack{a \leqq b \leqq x \\ ab \leqq x}} \frac{1}{a} + \sum_{\substack{b \leqq a \leqq x \\ ab \leqq x}} \frac{1}{a} - \sum_{\substack{a = b \leqq x \\ ab \leqq x}} \frac{1}{a} = S_1 + S_2 - S_3.$$

$$(6) \qquad S_1 = \sum_{a \leqq \sqrt{x}} \frac{1}{a} \sum_{a \leqq b \leqq \frac{x}{a}} 1 = \sum_{a \leqq \sqrt{x}} \frac{1}{a}\left(\left[\frac{x}{a}\right] + 1 - a\right)$$

$$(7) \qquad = -\sum_{a \leqq \sqrt{x}} \frac{1}{a}\psi\left(\frac{x}{a}\right) + \frac{1}{2}\sum_{a \leqq \sqrt{x}} \frac{1}{a} + x\sum_{a \leqq \sqrt{x}} \frac{1}{a^2} - [\sqrt{x}].$$

$$(8) \qquad \log[\sqrt{x}] = \log(\sqrt{x} - R(\sqrt{x})) = \frac{1}{2}\log x - \frac{R(\sqrt{x})}{\sqrt{x}} + \frac{B}{x}$$

$$(9) \qquad = \frac{1}{2}\log x + Bx^{-\frac{1}{2}}.$$

$$(10) \qquad \sum_{a \leqq \sqrt{x}} \frac{1}{a} = \log[\sqrt{x}] + \gamma + Bx^{-\frac{1}{2}} = \frac{1}{2}\log x + \gamma + Bx^{-\frac{1}{2}} \quad (9).$$

$$(11) \qquad q = [\sqrt{x}] + 1.$$

$$(11) \qquad \sum_{a=q}^{\infty} \frac{1}{a^2} = \sum_{a=q}^{\infty}\left(\frac{1}{a} - \frac{1}{a+1}\right) + \frac{1}{2}\sum_{a=q}^{\infty}\left(\frac{1}{a^2} - \frac{1}{(a+1)^2}\right) + \frac{1}{2}\sum_{a=q}^{\infty} \frac{1}{a^2(a+1)^2}$$

$$(12) \qquad = \frac{1}{q} + \frac{1}{2q^2} + \frac{B}{q^3}.$$

$$(13) \qquad \frac{1}{q} = \frac{1}{\sqrt{x} - R(\sqrt{x}) + 1} = \frac{1}{\sqrt{x}} + \frac{R(\sqrt{x}) - 1}{x} + Bx^{-\frac{3}{2}} \quad (11),$$

$$(14) \qquad \frac{1}{2q^2} = \frac{1}{2x} + Bx^{-\frac{3}{2}} \quad (13).$$

$$x\sum_{a > \sqrt{x}} \frac{1}{a^2} = \sqrt{x} + R(\sqrt{x}) - 1 + \frac{1}{2} + Bx^{-\frac{1}{2}} \quad (11 \text{ bis } 14)$$

$$(15) \qquad = 2\sqrt{x} - [\sqrt{x}] - \frac{1}{2} + Bx^{-\frac{1}{2}}.$$

$$x\sum_{a \leqq \sqrt{x}} \frac{1}{a^2} = \frac{\pi^2}{6}x - x\sum_{a > \sqrt{x}} \frac{1}{a^2}$$

$$(16) \qquad = \frac{\pi^2}{6}x - 2\sqrt{x} + [\sqrt{x}] + \frac{1}{2} + Bx^{-\frac{1}{2}} \quad (15).$$

(7, 10, 16) $$S_1 = -\sum_{a \leqq \sqrt{x}} \frac{1}{a}\psi\left(\frac{x}{a}\right) + \frac{1}{4}\log x + \frac{1}{2}\gamma + \frac{\pi^2}{6}x - 2\sqrt{x} + [\sqrt{x}] + \frac{1}{2} - [\sqrt{x}] + Bx^{-\frac{1}{2}}$$

(17) $$= \frac{\pi^2}{6}x - 2\sqrt{x} + \frac{1}{4}\log x + \frac{1}{2} + \frac{1}{2}\gamma - \sum_{a \leqq \sqrt{x}} \frac{1}{a}\psi\left(\frac{x}{a}\right) + Bx^{-\frac{1}{2}}.$$

(18) $$S_2 = \sum_{b \leqq \sqrt{x}} \sum_{b \leqq a \leqq \frac{x}{b}} \frac{1}{a} = \sum_{b \leqq \sqrt{x}} \sum_{a \leqq \frac{x}{b}}' \frac{1}{a} - \sum_{b \leqq \sqrt{x}} \sum_{a < b}' \frac{1}{a} \qquad (6)$$

(19) $$= S_{21} - S_{22}.$$

(20) $$\sum_{a \leqq \frac{x}{b}}' \frac{1}{a} = \log\left[\frac{x}{b}\right] + \gamma + \frac{1}{2}\left[\frac{x}{b}\right]^{-1} + Bb^2x^{-2}.$$

$$\log\left[\frac{x}{b}\right] = \log\left(\frac{x}{b} - R\left(\frac{x}{b}\right)\right) = \log\frac{x}{b} - \frac{b}{x}R\left(\frac{x}{b}\right) + Bb^2x^{-2}$$

(21) $$= \log x - \log b - \frac{b}{x}\psi\left(\frac{x}{b}\right) - \frac{b}{2x} + Bb^2x^{-2}.$$

(22) $$\left[\frac{x}{b}\right]^{-1} = \left(\frac{x}{b} - R\left(\frac{x}{b}\right)\right)^{-1} = \frac{b}{x} + Bb^2x^{-2}.$$

(23) $$\sum_{a \leqq \frac{x}{b}}' \frac{1}{a} = \log x - \log b - \frac{b}{x}\psi\left(\frac{x}{b}\right) + \gamma + Bb^2x^{-2} \quad (20, 21, 22).$$

$$\sum_{b \leqq \sqrt{x}} \log b = \left([\sqrt{x}] + \frac{1}{2}\right)\log[\sqrt{x}] - [\sqrt{x}] + \frac{1}{2}\log 2\pi + Bx^{-\frac{1}{2}}$$

$$= (\sqrt{x} - \psi(\sqrt{x}))\left(\frac{1}{2}\log x - \frac{R(\sqrt{x})}{\sqrt{x}} + Bx^{-1}\right) - \sqrt{x} + R(\sqrt{x}) + \frac{1}{2}\log 2\pi + Bx^{-\frac{1}{2}} \qquad (8)$$

(24) $$= \frac{1}{2}\left([\sqrt{x}] + \frac{1}{2}\right)\log x - \sqrt{x} + \frac{1}{2}\log 2\pi + Bx^{-\frac{1}{2}}.$$

$$S_{21} = [\sqrt{x}](\log x + \gamma) - \frac{1}{2}\left([\sqrt{x}] + \frac{1}{2}\right)\log x + \sqrt{x} - \frac{1}{2}\log 2\pi - \frac{1}{x}\sum_{b \leqq \sqrt{x}} b\psi\left(\frac{x}{b}\right) + Bx^{-\frac{1}{2}} \qquad (18, 19, 23, 24)$$

(25) $$= \frac{1}{2}[\sqrt{x}]\log x - \frac{1}{4}\log x + \gamma[\sqrt{x}] + \sqrt{x} - \frac{1}{2}\log 2\pi - \frac{1}{x}\sum_{b \leqq \sqrt{x}} b\psi\left(\frac{x}{b}\right) + Bx^{-\frac{1}{2}}.$$

(26) $$S_{22} = \sum_{b \leqq \sqrt{x}} \sum_{a \leqq b}' \frac{1}{a} - \sum_{b \leqq \sqrt{x}} \frac{1}{b} \qquad (18, 19).$$

$$\sum_{b\leqq\sqrt{x}}\sum_{a\leqq b}\frac{1}{a}=\sum_{a\leqq\sqrt{x}}\frac{1}{a}\sum_{a\leqq b\leqq\sqrt{x}}1=\sum_{a\leqq\sqrt{x}}\frac{1}{a}([\sqrt{x}]+1-a)$$

$$(27)\qquad =([\sqrt{x}]+1)\sum_{a\leqq\sqrt{x}}\frac{1}{a}-[\sqrt{x}].$$

$$S_{23}=[\sqrt{x}]\Big(\sum_{a\leqq\sqrt{x}}\frac{1}{a}-1\Big) \qquad (26,27)$$

$$=[\sqrt{x}]\Big(\log[\sqrt{x}]+\gamma-1+\frac{1}{2[\sqrt{x}]}+Bx^{-1}\Big)$$

$$=[\sqrt{x}]\log[\sqrt{x}]+\gamma[\sqrt{x}]-[\sqrt{x}]+\tfrac{1}{2}+Bx^{-\frac{1}{2}}$$

$$=\tfrac{1}{2}[\sqrt{x}]\log x+\gamma[\sqrt{x}]-\sqrt{x}+\tfrac{1}{2}+\frac{R^2(\sqrt{x})}{\sqrt{x}}+Bx^{-\frac{1}{2}} \qquad (8)$$

$$(28)\qquad =\tfrac{1}{2}[\sqrt{x}]\log x+\gamma[\sqrt{x}]-\sqrt{x}+\tfrac{1}{2}+Bx^{-\frac{1}{2}}.$$

$$(29)\qquad S_2=2\sqrt{x}-\tfrac{1}{4}\log x-\tfrac{1}{2}-\tfrac{1}{2}\log 2\pi-\frac{1}{x}\sum_{b\leqq\sqrt{x}}b\,\psi\Big(\frac{x}{b}\Big)+Bx^{-\frac{1}{2}}$$

(18, 19, 25, 28).

$$(30)\qquad S_3=\sum_{a\leqq\sqrt{x}}\frac{1}{a}=\tfrac{1}{2}\log x+\gamma+Bx^{-\frac{1}{2}} \qquad \textbf{(6, 10).}$$

$$(31)\quad S_0(x)=\frac{\pi^2}{6}x-\tfrac{1}{2}\log x-\tfrac{1}{2}(\gamma+\log 2\pi)-\sum_{a\leqq\sqrt{x}}\frac{1}{a}\psi\Big(\frac{x}{a}\Big)$$

$$-\frac{1}{x}\sum_{b\leqq\sqrt{x}}b\,\psi\Big(\frac{x}{b}\Big)+Bx^{-\frac{1}{2}} \qquad \textbf{(6, 17, 29, 30).}$$

Aus (1) und (31) folgt (5).

Bis auf weiteres sei jetzt $x\geqq 2$.

Hilfssatz 2: *Mit*

$$(32)\qquad c=\mathrm{Max}\,(a^2;b^2)$$

ist

$$(33)\quad \int_0^x T_0^2(u)\,du=\tfrac{1}{4}(\gamma+\log 2\pi)^2x+\sum_{a,b\leqq\sqrt{x}}\frac{1}{ab}\int_c^x\psi\Big(\frac{u}{a}\Big)\psi\Big(\frac{u}{b}\Big)\,du$$

$$+2\sum_{a,b\leqq\sqrt{x}}\frac{b}{a}\int_c^x\psi\Big(\frac{u}{a}\Big)\psi\Big(\frac{u}{b}\Big)\frac{du}{u}+\sum_{a,b\leqq\sqrt{x}}ab\int_c^x\psi\Big(\frac{u}{a}\Big)\psi\Big(\frac{u}{b}\Big)\frac{du}{u^2}+Bx^{\frac{1}{2}}\log x.$$

Beweis:

$$(1)\qquad \int_0^x T_0^2(u)\,du = \int_1^x T_0^2(u)\,du + B$$

$$(5)\quad = \tfrac{1}{4}(\gamma+\log 2\pi)^2 x + \int_1^x \sum_{a\leqq\sqrt{u}} \frac{1}{a}\psi\left(\frac{u}{a}\right) \sum_{b\leqq\sqrt{u}} \frac{1}{b}\psi\left(\frac{u}{b}\right) du$$

$$+ 2\int_1^x \sum_{a\leqq\sqrt{u}} \frac{1}{a}\psi\left(\frac{u}{a}\right) \sum_{b\leqq\sqrt{u}} b\psi\left(\frac{u}{b}\right)\frac{du}{u} + \int_1^x \sum_{a\leqq\sqrt{u}} a\psi\left(\frac{u}{a}\right) \sum_{b\leqq\sqrt{u}} b\psi\left(\frac{u}{b}\right)\frac{du}{u^2}$$

$$+ (\gamma+\log 2\pi)\int_1^x \sum_{a\leqq\sqrt{u}} \frac{1}{a}\psi\left(\frac{u}{a}\right) du + (\gamma+\log 2\pi)\int_1^x \sum_{b\leqq\sqrt{u}} b\psi\left(\frac{u}{b}\right)\frac{du}{u}$$

$$+ B + B\int_1^x u^{-\frac{1}{2}}\left(1 + \sum_{a\leqq\sqrt{u}} \frac{1}{a} + \frac{1}{u}\sum_{b\leqq\sqrt{u}} b\right) du + B\log x$$

$$(32)\quad = \tfrac{1}{4}(\gamma+\log 2\pi)^2 x + \sum_{a,b\leqq\sqrt{x}} \frac{1}{ab}\int_c^x \psi\left(\frac{u}{a}\right)\psi\left(\frac{u}{b}\right) du$$

$$+ 2\sum_{a,b\leqq\sqrt{x}} \frac{b}{a}\int_c^x \psi\left(\frac{u}{a}\right)\psi\left(\frac{u}{b}\right)\frac{du}{u} + \sum_{a,b\leqq\sqrt{x}} ab\int_c^x \psi\left(\frac{u}{a}\right)\psi\left(\frac{u}{b}\right)\frac{du}{u^2}$$

$$(34)\qquad + (\gamma+\log 2\pi)\sum_{a\leqq\sqrt{x}} \frac{1}{a}\int_{a^2}^x \psi\left(\frac{u}{a}\right) du + (\gamma+\log 2\pi)\sum_{b\leqq\sqrt{x}} b\int_{b^2}^x \psi\left(\frac{u}{b}\right)\frac{du}{u}$$

$$+ Bx^{\frac{1}{2}}\log x.$$

Über jedes endliche Intervall integriert, ist

$$(35)\qquad \int \psi\left(\frac{u}{a}\right) du = a\int \psi(u)\,du = Ba.$$

$$(36)\qquad \int_{a^2}^x \psi\left(\frac{u}{a}\right) du = Ba, \qquad \int_{b^2}^x \psi\left(\frac{u}{b}\right)\frac{du}{u} = B\frac{b}{b^2} = \frac{B}{b} \qquad (35),$$

$$(37)\qquad \sum_{a\leqq\sqrt{x}} \frac{1}{a}\int_{a^2}^x \psi\left(\frac{u}{a}\right) du = B\sqrt{x}, \qquad \sum_{b\leqq\sqrt{x}} b\int_{b^2}^x \psi\left(\frac{u}{b}\right)\frac{du}{u} = B\sqrt{x} \qquad (36).$$

Aus (34) und (37) folgt (33).

Hilfssatz 3:

$$(38)\qquad \sum_{\substack{m,n=1\\ mb=na}}^{\infty} \frac{1}{mn} = \frac{\pi^2}{6}\frac{(a,b)^2}{ab}.$$

Beweis:

$$mb = na; \quad m\frac{b}{(a,b)} = n\frac{a}{(a,b)}; \quad m = \mu\frac{a}{(a,b)}, \quad n = \nu\frac{b}{(a,b)};$$

$$\mu\frac{ab}{(a,b)^2} = \nu\frac{ab}{(a,b)^2}; \quad \mu = \nu; \quad m = \mu\frac{a}{(a,b)}, \quad n = \mu\frac{b}{(a,b)};$$

$$\sum_{\substack{m,n=1\\ mb=na}}^{\infty}\frac{1}{mn} = \frac{(a,b)^2}{ab}\sum_{\mu=1}^{\infty}\frac{1}{\mu^2} = \frac{\pi^2(a,b)^2}{6ab}.$$

Hilfssatz 4:

$$\sum_{\substack{a,b=1\\ (a,b)=1}}^{\infty}\frac{1}{a^2b^2} = \frac{5}{2}. \tag{39}$$

Beweis:

$$\frac{\pi^4}{36} = \sum_{a,b=1}^{\infty}\frac{1}{a^2b^2} = \sum_{d=1}^{\infty}\sum_{\substack{a,b=1\\ (a,b)=d}}^{\infty}\frac{1}{a^2b^2} = \sum_{d=1}^{\infty}\sum_{\substack{a,b=1\\ (a,b)=1}}^{\infty}\frac{1}{(da)^2(db)^2}$$

$$= \sum_{d=1}^{\infty}\frac{1}{d^4}\sum_{\substack{a,b=1\\ (a,b)=1}}^{\infty}\frac{1}{a^2b^2} = \frac{\pi^4}{90}\sum_{\substack{a,b=1\\ (a,b)=1}}^{\infty}\frac{1}{a^2b^2}.$$

Hilfssatz 5:

$$J_1 = \frac{1}{2\pi^2}\sum_{a,b\leqq\sqrt{x}}\frac{x-c}{ab}\sum_{\substack{m,n=1\\ mb=na}}^{\infty}\frac{1}{mn} \tag{40}$$

$$= \frac{5\pi^2}{144}x + Bx^{\frac{1}{2}}\log x. \tag{41}$$

Beweis:

$$J_1 = \frac{1}{12}\sum_{a,b\leqq\sqrt{x}}\frac{(x-c)(a,b)^2}{a^2b^2} \qquad (40,$$

$$= \frac{1}{12}\sum_{a\leqq b\leqq\sqrt{x}}\frac{(x-b^2)(a,b)^2}{a^2b^2} + \frac{1}{12}\sum_{b\leqq a\leqq\sqrt{x}}\frac{(x-a^2)(a,b)^2}{a^2b^2} - \frac{1}{12}\sum_{a=b\leqq\sqrt{x}}\frac{(x-a^2)(a,b)^2}{a^2b^2} \qquad (3$$

$$= \frac{1}{6}\sum_{a\leqq b\leqq\sqrt{x}}\frac{(x-b^2)(a,b)^2}{a^2b^2} - \frac{1}{12}\sum_{a\leqq\sqrt{x}}\frac{x-a^2}{a^2}. \tag{42}$$

$$\frac{1}{12}\sum_{a\leqq\sqrt{x}}{}'\frac{x-a^2}{a^2} = \frac{\pi^2}{72}x - \frac{x}{12}\sum_{a>\sqrt{x}}\frac{1}{a^2} + B\sqrt{x} = \frac{\pi^2}{72}x + B\sqrt{x}. \tag{43}$$

$$\sum_{a\leqq b\leqq\sqrt{x}}\frac{(x-b^2)(a,b)^2}{a^2b^2} = \sum_{d\leqq\sqrt{x}}d^2\sum_{\substack{a\leqq b\leqq\sqrt{x}\\ (a,b)=d}}\frac{x-b^2}{a^2b^2} = \sum_{d\leqq\sqrt{x}}d^2\sum_{\substack{a\leqq b\leqq\frac{\sqrt{x}}{d}\\ (a,b)=1}}\frac{x-d^2b^2}{d^4a^2b^2}$$

$$= x\sum_{d\leqq\sqrt{x}}\frac{1}{d^2}\sum_{b\leqq\frac{\sqrt{x}}{d}}\frac{1}{b^2}\sum_{\substack{a\leqq b\\ (a,b)=1}}\frac{1}{a^2} - \sum_{d\leqq\sqrt{x}}\sum_{b\leqq\frac{\sqrt{x}}{d}}\sum_{\substack{a\leqq b\\ (a,b)=1}}\frac{1}{a^2}. \tag{44}$$

(45) $$\sum_{d\leqq\sqrt{x}}\sum_{b\leqq\frac{\sqrt{x}}{d}}\sum_{\substack{a\leqq b\\(a,b)=1}}\frac{1}{a^2}=B\sum_{d\leqq\sqrt{x}}\sum_{b\leqq\frac{\sqrt{x}}{d}}1=Bx^{\frac{1}{2}}\log x.$$

(46) $$\sum_{d\leqq\sqrt{x}}\frac{1}{d^2}\sum_{b>\frac{\sqrt{x}}{d}}\frac{1}{b^2}\sum_{\substack{a\leqq b\\(a,b)=1}}\frac{1}{a^2}=B\sum_{d\leqq\sqrt{x}}\frac{1}{d^2}\frac{d}{\sqrt{x}}=Bx^{-\frac{1}{2}}\log x.$$

(47) $$\sum_{d>\sqrt{x}}\frac{1}{d^2}\sum_{b=1}^{\infty}\frac{1}{b^2}\sum_{\substack{a\leqq b\\(a,b)=1}}\frac{1}{a^2}=B\sum_{d>\sqrt{x}}\frac{1}{d^2}=Bx^{-\frac{1}{2}}.$$

$$J_{\mathrm{I}}=\frac{x}{6}\sum_{d=1}^{\infty}\frac{1}{d^2}\sum_{b=1}^{\infty}\frac{1}{b^2}\sum_{\substack{a\leqq b\\(a,b)=1}}\frac{1}{a^2}-\frac{\pi^2}{72}x+Bx^{\frac{1}{2}}\log x \qquad (42 \text{ bis } 47)$$

(48) $$=\frac{\pi^2}{36}x\sum_{b=1}^{\infty}\frac{1}{b^2}\sum_{\substack{a\leqq b\\(a,b)=1}}\frac{1}{a^2}-\frac{\pi^2}{72}x+Bx^{\frac{1}{2}}\log x.$$

(49) $$2\sum_{b=1}^{\infty}\frac{1}{b^2}\sum_{\substack{a\leqq b\\(a,b)=1}}\frac{1}{a^2}=\sum_{\substack{a,b=1\\(a,b)=1}}^{\infty}\frac{1}{a^2b^2}+\sum_{\substack{a,b=1\\a=b\\(a,b)=1}}^{\infty}\frac{1}{a^2b^2}=\frac{5}{2}+1=\frac{7}{2} \qquad (39).$$

Aus (48) und (49) folgt (41).

Hilfssatz 6:

(50) $$J_{\mathrm{II}}=\sum_{a\leqq b\leqq\sqrt{x}}\sum_{\substack{m,n=1\\mb\neq na}}^{\infty}\frac{1}{mn\,|mb-na|}$$

(51) $$=Bx^{\frac{1}{2}}\log x.$$

Beweis:

(52) $$J_{\mathrm{II}}=\sum_{a\leqq b\leqq\sqrt{x}}\left(\sum_{\substack{m,n=1\\na\leqq\frac{1}{2}mb}}^{\infty}+\sum_{\substack{m,n=1\\na\geqq 2mb}}^{\infty}+\sum_{\substack{m,n=1\\\frac{1}{2}mb<na<mb}}^{\infty}+\sum_{\substack{m,n=1\\mb<na<2mb}}^{\infty}\right)\frac{1}{mn\,|mb-na|} \qquad (50)$$

(53) $$=S_1'+S_2'+S_3'+S_4'.$$

S_1' ist für S_2' typisch, ebenso S_3' für S_4'. Daher gebe ich für die folgenden Formeln (55) und (68) keine Begründung an.

(52, 53) $$S_1'=B\sum_{a\leqq b\leqq\sqrt{x}}\sum_{\substack{m,n=1\\na\leqq\frac{1}{2}mb}}^{\infty}\frac{1}{mn\,mb}=B\sum_{a\leqq b\leqq\sqrt{x}}\frac{1}{b}\sum_{m=1}^{\infty}\frac{1}{m^2}\sum_{n\leqq\frac{1}{2}\frac{mb}{a}}\frac{1}{n}$$

(54) $$=B\sum_{a\leqq b\leqq\sqrt{x}}\frac{1}{b}\sum_{m=1}^{\infty}\frac{1}{m^2}(\log m+\log x)=B\log x\sum_{b\leqq\sqrt{x}}\frac{1}{b}\sum_{a\leqq b}1=Bx^{\frac{1}{2}}\log x.$$

(55) $$S_2'=Bx^{\frac{1}{2}}\log x.$$

$$(52,53)\quad S_3' = B\sum_{a\leqq b\leqq\sqrt{x}}\sum_{\substack{m,\,n=1\\ \frac{1}{2}mb<na<mb}}^{\infty}\frac{a}{m^2b(mb-na)} = B\sum_{a\leqq b\leqq\sqrt{x}}\frac{1}{b}\sum_{m=1}^{\infty}\frac{1}{m^2}\sum_{n<\frac{mb}{a}}\frac{1}{\frac{mb}{a}-n}$$

$$= B\sum_{a\leqq b\leqq\sqrt{x}}\frac{1}{b}\sum_{m=1}^{\infty}\frac{1}{m^2}\left(\sum_{\frac{mb}{a}-1<n<\frac{mb}{a}}+\sum_{n\leqq\frac{mb}{a}-1}\right)\frac{1}{\frac{mb}{a}-n}$$

$$= B\sum_{a\leqq b\leqq\sqrt{x}}\frac{1}{b}\sum_{\substack{m=1\\ a\nmid mb}}^{\infty}\frac{1}{m^2R\left(\frac{mb}{a}\right)}+B\sum_{a\leqq b\leqq\sqrt{x}}\frac{1}{b}\sum_{m=1}^{\infty}\frac{1}{m^2}(\log m+\log x)$$

$$(56)\qquad = B\sum_{m=1}^{\infty}\frac{1}{m^2}\sum_{a\leqq\sqrt{x}}\sum_{\substack{a\leqq b\leqq\sqrt{x}\\ a\nmid mb}}\frac{1}{bR\left(\frac{mb}{a}\right)}+Bx^{\frac{1}{2}}\log x \qquad (54).$$

$$(57)\qquad S_3'' = \sum_{\substack{a\leqq b\leqq\sqrt{x}\\ a\nmid mb}}\frac{1}{bR\left(\frac{mb}{a}\right)}.$$

$$(58)\qquad a=\alpha(a,m),\qquad m=\mu(a,m).$$

$$(59)\qquad S_3'' = \sum_{\substack{a\leqq b\leqq\sqrt{x}\\ \alpha\nmid b}}\frac{1}{bR\left(\frac{\mu b}{\alpha}\right)} \qquad (57,58).$$

$$(60)\qquad \tau(b)=\begin{cases} R^{-1}\left(\frac{\mu b}{\alpha}\right) & \text{für } \alpha\nmid b\\ 0 & \text{für } \alpha\mid b.\end{cases}$$

$$(61)\qquad \sum_{b=a}^{N}\tau(b)=\varrho(N)\ \text{ für }\ N\geqq a,\qquad \varrho(a-1)=0.$$

$$S_3'' = \sum_{a\leqq b\leqq\sqrt{x}}\frac{\tau(b)}{b} = \sum_{a\leqq b\leqq\sqrt{x}}\varrho(b)\left(\frac{1}{b}-\frac{1}{b+1}\right)+\frac{\varrho([\sqrt{x}])}{[\sqrt{x}]+1} \qquad (59,60,61)$$

$$(62)\qquad = \sum_{a\leqq b\leqq\sqrt{x}}\frac{\varrho(b)}{b(b+1)}+\frac{\varrho([\sqrt{x}])}{[\sqrt{x}]+1}.$$

$$\sum_{\substack{b=0\\ (\mathrm{mod}\,\alpha)}}^{\alpha-1}\tau(b)=\sum_{b=1}^{\alpha-1}R^{-1}\left(\frac{\mu b}{\alpha}\right)=\sum_{b=1}^{\alpha-1}R^{-1}\left(\frac{b}{\alpha}\right) \qquad (60,58)$$

$$(63)\qquad = \alpha\sum_{b=1}^{\alpha-1}\frac{1}{b}=B\alpha\log x \qquad (57,58).$$

$$(64)\qquad (k-1)\alpha<b\leqq k\alpha \qquad (k \text{ ganzzahlig}).$$

$$\varrho(b) \leqq \sum_{r=a}^{k\alpha} \tau(r) \quad (61, 64) \quad = B\frac{k\alpha}{\alpha}\alpha \log x = B k\alpha \log x \tag{63}$$

$$= B(b+\alpha)\log x = B(b+a)\log x = B b \log x \quad (64, 58, 57). \tag{65}$$

$$S_3'' = B \log x \sum_{a \leqq b \leqq \sqrt{x}} \frac{1}{b} + B \log x \tag{66}$$

(62, 65).

$$S_3' = B \log x \sum_{m=1}^{\infty} \frac{1}{m^2} \sum_{a \leqq \sqrt{x}} \sum_{a \leqq b \leqq \sqrt{x}} \frac{1}{b} + B \log x \sum_{m=1}^{\infty} \frac{1}{m^2} \sum_{a \leqq \sqrt{x}} 1$$

$$+ B x^{\frac{1}{2}} \log x \qquad (56, 57, 66)$$

$$= B \log x \sum_{b \leqq \sqrt{x}} \frac{1}{b} \sum_{a \leqq b} 1 + B x^{\frac{1}{2}} \log x = B x^{\frac{1}{2}} \log x. \tag{67}$$

$$S_4' = B x^{\frac{1}{2}} \log x. \tag{68}$$

Aus (52), (53), (54), (55), (67) und (68) folgt (51).

Hilfssatz 7:

$$\sum_{a,b \leqq \sqrt{x}} \frac{1}{ab} \sum_{\substack{m,n=1 \\ mb \neq na}}^{\infty} \frac{1}{mn} \left| \int_c^x \cos 2\pi u \left(\frac{m}{a} - \frac{n}{b}\right) du \right| = B x^{\frac{1}{2}} \log x, \tag{69}$$

$$\sum_{a,b \leqq \sqrt{x}} \frac{b}{a} \sum_{\substack{m,n=1 \\ mb \neq na}}^{\infty} \frac{1}{mn} \left| \int_c^x \cos 2\pi u \left(\frac{m}{a} - \frac{n}{b}\right) \frac{du}{u} \right| = B x^{\frac{1}{2}} \log x, \tag{70}$$

$$\sum_{a,b \leqq \sqrt{x}} ab \sum_{\substack{m,n=1 \\ mb \neq na}}^{\infty} \frac{1}{mn} \left| \int_c^x \cos 2\pi u \left(\frac{m}{a} - \frac{n}{b}\right) \frac{du}{u^2} \right| = B x^{\frac{1}{2}} \log x. \tag{71}$$

Beweis. Die linke Seite von (69) ist

$$= B \sum_{a,b \leqq \sqrt{x}} \frac{1}{ab} \sum_{\substack{m,n=1 \\ mb \neq na}}^{\infty} \frac{1}{mn \left| \frac{m}{a} - \frac{n}{b} \right|} = B \sum_{a,b \leqq \sqrt{x}} \sum_{\substack{m,n=1 \\ mb \neq na}}^{\infty} \frac{1}{mn\,|mb - na|} \tag{72}$$

$$= B \sum_{a \leqq b \leqq \sqrt{x}} \sum_{\substack{m,n=1 \\ mb \neq na}}^{\infty} \frac{1}{mn\,|mb - na|} = B x^{\frac{1}{2}} \log x \qquad (50, 51). \tag{73}$$

Die linke Seite von (70) ist

$$= B \sum_{a,b \leqq \sqrt{x}} \frac{b}{ac} \sum_{\substack{m,n=1 \\ mb \neq na}}^{\infty} \frac{1}{mn \left| \frac{m}{a} - \frac{n}{b} \right|} = B \sum_{a,b \leqq \sqrt{x}} \frac{1}{ab} \sum_{\substack{m,n=1 \\ mb \neq na}}^{\infty} \frac{1}{mn \left| \frac{m}{a} - \frac{n}{b} \right|} \tag{32}$$

$$= B x^{\frac{1}{2}} \log x \qquad (72, 73).$$

Die linke Seite von (71) ist

$$= B \sum_{a,b \leqq \sqrt{x}} \frac{ab}{c^2} \sum_{\substack{m,n=1 \\ mb \neq na}}^{\infty} \frac{1}{mn \left| \frac{m}{a} - \frac{n}{b} \right|} = B \sum_{a,b \leqq \sqrt{x}} \frac{1}{ab} \sum_{\substack{m,n=1 \\ mb \neq na}}^{\infty} \frac{1}{mn \left| \frac{m}{a} - \frac{n}{b} \right|} \quad (32)$$

$$= B x^{\frac{1}{2}} \log x \quad (72, 73).$$

Hilfssatz 8:

$$(74) \qquad \sum_{a,b \leqq \sqrt{x}} \frac{1}{ab} \sum_{m,n=1}^{\infty} \frac{1}{mn} \left| \int_c^x \cos 2\pi u \left(\frac{m}{a} + \frac{n}{b} \right) du \right| = B x^{\frac{1}{2}},$$

$$(75) \qquad \sum_{a,b \leqq \sqrt{x}} \frac{b}{a} \sum_{m,n=1}^{\infty} \frac{1}{mn} \left| \int_c^x \cos 2\pi u \left(\frac{m}{a} + \frac{n}{b} \right) \frac{du}{u} \right| = B x^{\frac{1}{2}}.$$

$$(76) \qquad \sum_{a,b \leqq \sqrt{x}} ab \sum_{m,n=1}^{\infty} \frac{1}{mn} \left| \int_c^x \cos 2\pi u \left(\frac{m}{a} + \frac{n}{b} \right) \frac{du}{u^2} \right| = B x^{\frac{1}{2}}.$$

Beweis: Die linke Seite von (74) ist

$$(77) \quad = B \sum_{a,b \leqq \sqrt{x}} \frac{1}{ab} \sum_{m,n=1}^{\infty} \frac{1}{mn \left(\frac{m}{a} + \frac{n}{b} \right)} = B \sum_{a,b \leqq \sqrt{x}} \frac{1}{ab} \sum_{m,n=1}^{\infty} \frac{1}{mn \sqrt{\frac{m}{a} \frac{n}{b}}}$$

$$(78) \quad = B \sum_{a,b \leqq \sqrt{x}} \frac{1}{\sqrt{ab}} = B x^{\frac{1}{2}}.$$

Die linke Seite von (75) ist

$$= B \sum_{a,b \leqq \sqrt{x}} \frac{b}{ac} \sum_{m,n=1}^{\infty} \frac{1}{mn \left(\frac{m}{a} + \frac{n}{b} \right)} = B \sum_{a,b \leqq \sqrt{x}} \frac{1}{ab} \sum_{m,n=1}^{\infty} \frac{1}{mn \left(\frac{m}{a} + \frac{n}{b} \right)} \quad (32)$$

$$= B x^{\frac{1}{2}} \quad (77, 78).$$

Die linke Seite von (76) ist

$$= B \sum_{a,b \leqq \sqrt{x}} \frac{ab}{c^2} \sum_{m,n=1}^{\infty} \frac{1}{mn \left(\frac{m}{a} + \frac{n}{b} \right)} = B \sum_{a,b \leqq \sqrt{x}} \frac{1}{ab} \sum_{m,n=1}^{\infty} \frac{1}{mn \left(\frac{m}{a} + \frac{n}{b} \right)} \quad (32)$$

$$= B x^{\frac{1}{2}} \quad (77, 78).$$

Hilfssatz 9:

$$(79) \qquad \sum_{a,b \leqq \sqrt{x}} \frac{b}{a} \log \frac{x}{c} \sum_{\substack{m,n=1 \\ mb = na}}^{\infty} \frac{1}{mn} = B x^{\frac{1}{2}} \log x,$$

$$(80) \qquad \sum_{a,b \leqq \sqrt{x}} \frac{ab}{c} \sum_{\substack{m,n=1 \\ mb = na}}^{\infty} \frac{1}{mn} = B x^{\frac{1}{2}} \log x.$$

Beweis von (79):

$$\sum_{a,b\leqq\sqrt{x}}\frac{b}{a}\log\frac{x}{c}\sum_{\substack{m,n=1\\ mb=na}}^{\infty}\frac{1}{mn}=B\sum_{a,b\leqq\sqrt{x}}\frac{(a,b)^2}{a^2}\log\frac{x}{c} \tag{38}$$

(32) $$=B\sum_{a\leqq b\leqq\sqrt{x}}\frac{(a,b)^2}{a^2}\log\frac{x}{b^2}=B\sum_{d\leqq\sqrt{x}}d^2\sum_{a\leqq b\leqq\frac{\sqrt{x}}{d}}\frac{1}{d^2a^2}\log\frac{x}{d^2b^2}$$

$$=B\sum_{d\leqq\sqrt{x}}\sum_{b\leqq\frac{\sqrt{x}}{d}}\log\frac{x}{d^2b^2}=B\sum_{d\leqq\sqrt{x}}\sum_{b\leqq\frac{\sqrt{x}}{d}}\int_1^{\frac{x}{d^2b^2}}\frac{dv}{v}$$

$$=B\int_1^x\frac{dv}{v}\sum_{d\leqq\sqrt{x}}\sum_{b\leqq\frac{1}{d}\sqrt{\frac{x}{v}}}1=Bx^{\frac{1}{2}}\int_1^x\frac{dv}{v^{\frac{1}{2}}}\sum_{d\leqq\sqrt{x}}\frac{1}{d}=Bx^{\frac{1}{2}}\log x.$$

Beweis von (80):

$$\sum_{a,b\leqq\sqrt{x}}\frac{ab}{c}\sum_{\substack{m,n=1\\ mb=na}}^{\infty}\frac{1}{mn}=B\sum_{a,b\leqq\sqrt{x}}\frac{(a,b)^2}{c} \tag{38}$$

(32) $$=B\sum_{a\leqq b\leqq\sqrt{x}}\frac{(a,b)^2}{a^2}=B\sum_{d\leqq\sqrt{x}}d^2\sum_{a\leqq b\leqq\frac{\sqrt{x}}{d}}\frac{1}{d^2a^2}=B\sum_{d\leqq\sqrt{x}}\sum_{b\leqq\frac{\sqrt{x}}{d}}1$$

$$=Bx^{\frac{1}{2}}\log x.$$

Hilfssatz 10: *Sind in dem Intervall $\alpha\leqq u\leqq\beta$ die quadratisch integrierbaren Funktionen $g(u)$, $h(u)$, $s_k(u)$, $t_k(u)$ $(k\geqq 1)$ gegeben, für die bei wachsendem k*

(81) $$\int_\alpha^\beta s_k^2(u)\,du,\qquad \int_\alpha^\beta t_k^2(u)\,du \quad \textit{beschränkt},$$

(82) $$\int_\alpha^\beta (g(u)-s_k(u))^2\,du\to 0,\qquad \int_\alpha^\beta (h(u)-t_k(u))^2\,du\to 0,$$

so ist

(83) $$\int_\alpha^\beta g(u)h(u)\,du=\lim_{k=\infty}\int_\alpha^\beta s_k(u)t_k(u)\,du.$$

Beweis:

$$\int_\alpha^\beta g(u)h(u)du-\int_\alpha^\beta s_k(u)t_k(u)\,du$$

$$=\int_\alpha^\beta (g(u)-s_k(u))t_k(u)\,du+\int_\alpha^\beta (h(u)-t_k(u))s_k(u)\,du$$

$$+\int_\alpha^\beta (g(u)-s_k(u))(h(u)-t_k(u))\,du,$$

und jedes der drei letzten Integrale strebt nach der Schwarzschen Ungleichung mit wachsendem k gegen Null (81, 82).

Hilfssatz 11:

$$(84)\qquad \sum_{a,b\leqq\sqrt{x}}\frac{1}{ab}\int_c^x \psi\left(\frac{u}{a}\right)\psi\left(\frac{u}{b}\right)du=\frac{5\pi^2}{144}x+Bx^{\frac{1}{2}}\log x.$$

Beweis: Mit $\alpha=c$, $\beta=x$, $g(u)=\psi\left(\frac{u}{a}\right)$, $h(u)=\psi\left(\frac{u}{b}\right)$,

$$(85)\qquad s(u)=-\frac{1}{\pi}\sum_{m=1}^{k}\frac{\sin 2\pi\frac{m}{a}u}{m},\qquad t_k(u)=-\frac{1}{\pi}\sum_{n=1}^{k}\frac{\sin 2\pi\frac{n}{b}u}{n}$$

sind die Voraussetzungen von Hilfssatz 10 erfüllt[3]). Daher liefern (83), (85), (40), (41), (69) und (74)

$$\sum_{a,b\leqq\sqrt{x}}\frac{1}{ab}\int_c^x \psi\left(\frac{u}{a}\right)\psi\left(\frac{u}{b}\right)du$$

$$=\lim_{k=\infty}\frac{1}{\pi^2}\sum_{a,b\leqq\sqrt{x}}\frac{1}{ab}\sum_{m,n=1}^{k}\frac{1}{mn}\int_c^x \sin 2\pi\frac{m}{a}u\sin 2\pi\frac{n}{b}u\,du$$

$$=\lim_{k=\infty}\frac{1}{2\pi^2}\sum_{a,b\leqq\sqrt{x}}\frac{1}{ab}\left(\sum_{\substack{m,n=1\\ mb=na}}^{k}\frac{1}{mn}\int_c^x du+\sum_{\substack{m,n=1\\ mb\neq na}}^{k}\frac{1}{mn}\int_c^x \cos 2\pi u\left(\frac{m}{a}-\frac{n}{b}\right)du\right.$$

$$\left.-\sum_{m,n=1}^{k}\frac{1}{mn}\int_c^x \cos 2\pi u\left(\frac{m}{a}+\frac{n}{b}\right)du\right)=\frac{5\pi^2}{144}x+Bx^{\frac{1}{2}}\log x.$$

Hilfssatz 12:

$$(86)\qquad \sum_{a,b\leqq\sqrt{x}}\frac{b}{a}\int_c^x \psi\left(\frac{u}{a}\right)\psi\left(\frac{u}{b}\right)\frac{du}{u}=Bx^{\frac{1}{2}}\log x.$$

Beweis: Mit $\alpha=c$, $\beta=x$, $g(u)=\frac{1}{\sqrt{u}}\psi\left(\frac{u}{a}\right)$, $h(u)=\frac{1}{\sqrt{u}}\psi\left(\frac{u}{b}\right)$,

$$(87)\qquad s_k(u)=-\frac{1}{\pi\sqrt{u}}\sum_{m=1}^{k}\frac{\sin 2\pi\frac{m}{a}u}{m},\qquad t_k(u)=-\frac{1}{\pi\sqrt{u}}\sum_{n=1}^{k}\frac{\sin 2\pi\frac{n}{b}u}{n}$$

sind die Voraussetzungen von Hilfssatz 10 erfüllt[3]). Daher liefern (83),

[3]) Vgl. hierzu etwa W. Rogosinski „Fouriersche Reihen" [Sammlung Göschen Nr. 1022 (1930)], Formeln (16), (40), (41).

(87), (79), (70) und (75)

$$\sum_{a,b\leqq\sqrt{x}}\frac{b}{a}\int_c^x \psi\left(\frac{u}{a}\right)\psi\left(\frac{u}{b}\right)\frac{du}{u}$$

$$=\lim_{k=\infty}\frac{1}{\pi^2}\sum_{a,b\leqq\sqrt{x}}\frac{b}{a}\sum_{m,n=1}^{k}\frac{1}{mn}\int_c^x \sin 2\pi\frac{m}{a}u\sin 2\pi\frac{n}{b}u\frac{du}{u}$$

$$=\lim_{k=\infty}\frac{1}{2\pi^2}\sum_{a,b\leqq\sqrt{x}}\frac{b}{a}\left(\sum_{\substack{m,n=1\\ mb=na}}^{k}\frac{1}{mn}\int_c^x\frac{du}{u}+\sum_{\substack{m,n=1\\ mb\neq na}}^{k}\frac{1}{mn}\int_c^x\cos 2\pi u\left(\frac{m}{a}-\frac{n}{b}\right)\frac{du}{u}\right.$$

$$\left.-\sum_{m,n=1}^{k}\frac{1}{mn}\int_c^x\cos 2\pi u\left(\frac{m}{a}+\frac{n}{b}\right)\frac{du}{u}\right)=Bx^{\frac{1}{2}}\log x.$$

Hilfssatz 13:

(88) $$\sum_{a,b\leqq\sqrt{x}}ab\int_c^x\psi\left(\frac{u}{a}\right)\psi\left(\frac{u}{b}\right)\frac{du}{u^2}=Bx^{\frac{1}{2}}\log x.$$

Beweis: Mit $\alpha=c$, $\beta=x$, $g(u)=\frac{1}{u}\psi\left(\frac{u}{a}\right)$, $h(u)=\frac{1}{u}\psi\left(\frac{u}{b}\right)$,

(89) $$s_k(u)=-\frac{1}{\pi u}\sum_{m=1}^{k}\frac{\sin 2\pi\frac{m}{a}u}{m},\qquad t_k(u)=-\frac{1}{\pi u}\sum_{n=1}^{k}\frac{\sin 2\pi\frac{n}{b}u}{n}$$

sind die Voraussetzungen von Hilfssatz 10 erfüllt [3]). Daher liefern (83), (89), (80), (71) und (76)

$$\sum_{a,b\leqq\sqrt{x}}ab\int_c^x\psi\left(\frac{u}{a}\right)\psi\left(\frac{u}{b}\right)\frac{du}{u^2}$$

$$=\lim_{k=\infty}\frac{1}{\pi^2}\sum_{a,b\leqq\sqrt{x}}ab\sum_{m,n=1}^{k}\frac{1}{mn}\int_c^x\sin 2\pi\frac{m}{a}u\sin 2\pi\frac{n}{b}u\frac{du}{u^2}$$

$$=\lim_{k=\infty}\frac{1}{2\pi^2}\sum_{a,b\leqq\sqrt{x}}ab\left(\sum_{\substack{m,n=1\\ mb=na}}^{k}\frac{1}{mn}\int_c^x\frac{du}{u^2}+\sum_{\substack{m,n=1\\ mb\neq na}}^{k}\frac{1}{mn}\int_c^x\cos 2\pi u\left(\frac{m}{a}-\frac{n}{b}\right)\frac{du}{u^2}\right.$$

$$\left.-\sum_{m,n=1}^{k}\frac{1}{mn}\int_c^x\cos 2\pi u\left(\frac{m}{a}+\frac{n}{b}\right)\frac{du}{u^2}\right)=Bx^{\frac{1}{2}}\log x.$$

Beweis von (I): (33), (84), (86), (88).

§ 2.
Beweis von (II).

Hilfssatz 14:[4])

$$T(x) - x\,T_0(x) = \frac{1}{2}(\gamma + \log 2\pi - 1)\,x + B\,x^{\frac{1}{2}} \qquad (x \geqq 1). \tag{90}$$

Beweis:

$$\int_1^x S_0(u)\,du = \frac{\pi^2}{12}x^2 - \frac{1}{2}x\log x + \frac{1}{2}x + \int_1^x T_0(u)\,du + B \tag{1}$$

$$= \frac{\pi^2}{12}x^2 - \frac{1}{2}x\log x + \frac{1}{2}x - \frac{1}{2}(\gamma + \log 2\pi)\,x$$

$$- \sum_{a \leqq \sqrt{x}} \frac{1}{a}\int_{a^2}^x \psi\left(\frac{u}{a}\right)du - \sum_{b \leqq \sqrt{x}} b\int_{b^2}^x \psi\left(\frac{u}{b}\right)\frac{du}{u} + B\,x^{\frac{1}{2}} \tag{5}$$

$$= \frac{\pi^2}{12}x^2 - \frac{1}{2}x\log x - \frac{1}{2}(\gamma + \log 2\pi - 1)\,x + B\,x^{\frac{1}{2}} \tag{91}$$ (37).

$$\int_1^x S_0(u)\,du = \int_1^x \sum_{n \leqq u} \sigma(n)\,du = \sum_{n \leqq x} \sigma(n)\int_n^x du = \sum_{n \leqq x}\sigma(n)\,(x-n)$$

$$= x\,S_0(x) - S(x) = \frac{\pi^2}{12}x^2 - \frac{1}{2}x\log x + x\,T_0(x) - T(x) \tag{92}$$ (1, 2).

Aus (91) und (92) folgt (90).

Es sei wieder $x \geqq 2$. Für den Rest dieses Paragraphen führe ich folgende Abkürzungen ein:

$$\lambda = \frac{1}{2}(\gamma + \log 2\pi - 1), \qquad \mu = \left(\lambda + \frac{1}{2}\right)^2 + \frac{5\pi^2}{144}, \tag{93}$$

$$\int_1^x T_0^2(u)\,du = T_1(x), \quad \int_1^x T_1(u)\,du = T_2(x), \quad \int_1^x T_2(u)\,du = T_3(x). \tag{94}$$

Beweis von (II):

$$\int_0^x T^2(u)\,du = \int_1^x T^2(u)\,du + B \tag{95}$$ (2).

---

[4]) Für die linke Seite von (90) sind folgende Restglieder bekannt: 1) $O\left(x^{\frac{3}{4}}\right)$ durch S. Wigert „Sur quelques fonctions arithmétiques" [Acta Mathematica **37** (1914), S. 113–140]; 2) $O\left(x^{\frac{1}{3}}\right)$ und $O\left(x^{\frac{3}{5}}\right)$ durch E. Landau „Besprechung von S. Wigert: Sur quelques fonctions arithmétiques" [Göttingische gelehrte Anzeigen **177** (1915), S. 377–414]; 3) $O\left(x^{\frac{1}{3}}\right)$ und $O\left(x^{\frac{7}{23}}\log^{\frac{12}{23}}x\right)$ durch E. Landau „Über einige zahlentheoretische Funktionen" [Göttinger Nachrichten (1924), S. 116–134]. Einen einfachen Beweis für $O\left(x^{\frac{1}{2}}\right)$ gab auch S. Ramanujan „On certain trigonometrical sums and their applications in the theory of numbers" [Transactions of the Cambridge Philosophical Society **22** (1918), S. 259–276 oder Collected Papers of Srinivasa Ramanujan (Cambridge 1927), S. 179–199].

$$\int_1^x T^2(u)\,du = \int_1^x (u\,T_0(u)+\lambda u)^2\,du + \int_1^x \left(T^2(u) - (u\,T_0(u)+\lambda u)^2\right) du$$

$$= \int_1^x u^2\,T_0^2(u)\,du + 2\lambda \int_1^x u^2\,T_0(u)\,du + \frac{\lambda^2 x^3}{3} \tag{96}$$

$$+ B\int_1^x |\,T(u)+u\,T_0(u)+\lambda u\,|\,|\,T(u)-u\,T_0(u)-\lambda u\,|\,du + B.$$

$$\int_1^x u^2\,T_0^2(u)\,du = u^2\,T_1(u)\Big|_1^x - 2\int_1^x u\,T_1(u)\,du \tag{94}$$

$$= x^2\,T_1(x) - 2u\,T_2(u)\Big|_1^x + 2\int_1^x T_2(u)\,du$$

$$= x^2\,T_1(x) - 2x\,T_2(x) + 2\,T_3(x). \tag{97}$$

$$T_1(x) = \mu x + B x^{\frac{1}{2}}\log x \qquad (94, 1, \text{I}, 93), \tag{98}$$

$$T_2(x) = \mu\frac{x^2}{2} + B x^{\frac{3}{2}}\log x \qquad (94, 98) \tag{99}$$

$$T_3(x) = \mu\frac{x^3}{6} + B x^{\frac{4}{2}}\log x \qquad (94, 99), \tag{100}$$

$$\int_1^x u^2\,T_0^2(u)\,du = \mu\frac{x^3}{3} + B x^{\frac{5}{2}}\log x \qquad (97 \text{ bis } 100). \tag{101}$$

$$-\int_1^x u^2\,T_0(u)\,du = \left(\lambda+\frac{1}{2}\right)\frac{x^3}{3} + \sum_{a\leqq\sqrt{x}}\frac{1}{a}\int_{a^2}^x \psi\left(\frac{u}{a}\right)u^2\,du$$

$$+ \sum_{b\leqq\sqrt{x}} b\int_{b^2}^x \psi\left(\frac{u}{b}\right)u\,du + B x^{\frac{5}{2}} \qquad (5, 93)$$

$$= \left(\lambda+\frac{1}{2}\right)\frac{x^3}{3} + B x^2 \sum_{a\leqq\sqrt{x}}\frac{1}{a}\,a + B x \sum_{b\leqq\sqrt{x}} b\,b + B x^{\frac{5}{2}} \qquad (35)$$

$$= \left(\lambda+\frac{1}{2}\right)\frac{x^3}{3} + B x^{\frac{5}{2}}. \tag{102}$$

$$\int_1^x |\,T(u)+u\,T_0(u)+\lambda u\,|\,|\,T(u)-u\,T_0(u)-\lambda u\,|\,du$$

$$= B\int_1^x u\log 2u\cdot u^{\frac{1}{2}}\,du \qquad (90, 93)$$

$$= B x^{\frac{5}{2}}\log x. \tag{103}$$

$$\int_0^x T^2(u)\,du = \left(\left(\lambda+\frac{1}{2}\right)^3 + \frac{5\pi^2}{144} - 2\lambda^2 - \lambda + \lambda^2\right)\frac{x^3}{3}$$

$$+ B x^{\frac{5}{2}}\log x \qquad (93, 95, 96, 101, 102, 103)$$

$$= \frac{36+5\pi^2}{432}x^3 + B x^{\frac{5}{2}}\log x. \tag{II}$$

§ 3.

Beweis von (III).

Es sei von jetzt ab dauernd $x \geqq 8$.

Hilfssatz 15:

$$\int_0^x \mathsf{P}_4^2(u)\,du = \frac{288+40\pi^2}{27}x^3 - 2^{11}\int_0^{\frac{x}{4}} T(4u)\,T(u)\,du + Bx^{\frac{5}{2}}\log x. \tag{104}$$

Beweis:

$$A_4(x) = 8S(x) - 32S\left(\frac{x}{4}\right) + 1,^{5}) \tag{105}$$

$$\mathsf{P}_4(x) = 8T(x) - 32T\left(\frac{x}{4}\right) + 1 \qquad (2, 3, 105). \tag{106}$$

$$\int_0^x \mathsf{P}_4^2(u)\,du = 2^6\int_0^x T^2(u)\,du + 2^{10}\int_0^x T^2\left(\frac{u}{4}\right)du - 2^9\int_0^x T(u)\,T\left(\frac{u}{4}\right)du + Bx^2\log x \qquad (106)$$

$$= 2^6\int_0^{\frac{x}{4}} T^2(u)\,du + 2^{12}\int_0^{\frac{x}{4}} T^2(u)\,du - 2^{11}\int_0^{\frac{x}{4}} T(4u)\,T(u)\,du + Bx^2\log x. \tag{107}$$

Aus (107) und (II) folgt (104).

Hilfssatz 16: *Mit*

$$r = \mathrm{Max}\left(\frac{a^2}{4}; b^2\right) \tag{108}$$

ist

$$\int_0^x T(4u)\,T(u)\,du = \frac{x^3}{3} + 4\sum_{\substack{a\leqq 2\sqrt{x}\\ b\leqq\sqrt{x}}}\frac{1}{ab}\int_r^x \psi\left(\frac{4u}{a}\right)\psi\left(\frac{u}{b}\right)u^2\,du + 4\sum_{\substack{a\leqq 2\sqrt{x}\\ b\leqq\sqrt{x}}}\frac{b}{a}\int_r^x \psi\left(\frac{4u}{a}\right)\psi\left(\frac{u}{b}\right)u\,du + \sum_{\substack{a\leqq 2\sqrt{x}\\ b\leqq\sqrt{x}}}\frac{a}{b}\int_r^x \psi\left(\frac{4u}{a}\right)\psi\left(\frac{u}{b}\right)u\,du + \sum_{\substack{a\leqq 2\sqrt{x}\\ b\leqq\sqrt{x}}} ab\int_r^x \psi\left(\frac{4u}{a}\right)\psi\left(\frac{u}{b}\right)du + Bx^{\frac{5}{2}}\log x. \tag{109}$$

Beweis:

$$-T(u) = \frac{u}{2} + u\sum_{a\leqq\sqrt{u}}\frac{1}{a}\psi\left(\frac{u}{a}\right) + \sum_{b\leqq\sqrt{u}} b\,\psi\left(\frac{u}{b}\right) + Bu^{\frac{1}{2}} \text{ für } u \geqq 1 \qquad (90, 5). \tag{110}$$

---

[5]) W. Sierpiński „Teorja liczb“ [Warszawa 1925, XII + 410 S.], S. 365 (76). Vgl. auch I, S. 69 (2. 3).

$$1) \quad \int_1^x \sum_{a \leqq \sqrt{u}} \frac{1}{a} \psi\left(\frac{u}{a}\right) u^2 du = \sum_{a \leqq \sqrt{x}} \frac{1}{a} \int_{a^2}^x \psi\left(\frac{u}{a}\right) u^2 du = B \sum_{a \leqq \sqrt{x}} \frac{1}{a} x^2 a = B x^{\frac{5}{2}} \quad (35),$$

$$2) \quad \int_1^x \sum_{b \leqq \sqrt{u}} b \psi\left(\frac{u}{b}\right) u\, du = \sum_{b \leqq \sqrt{x}} b \int_{b^2}^x \psi\left(\frac{u}{b}\right) u\, du = B \sum_{b \leqq \sqrt{x}} b x b = B x^{\frac{5}{2}} \quad (35),$$

$$3) \quad \int_1^x \sum_{a \leqq 2\sqrt{u}} \frac{1}{a} \psi\left(\frac{4u}{a}\right) u^2 du = \sum_{a \leqq 2\sqrt{x}} \frac{1}{a} \int_{\mathrm{Max}\left(\frac{a^2}{4};1\right)}^x \psi\left(\frac{4u}{a}\right) u^2 du = B \sum_{a \leqq 2\sqrt{x}} \frac{1}{a} x^2 a$$

$$= B x^{\frac{5}{2}} \quad (35),$$

$$4) \quad \int_1^x \sum_{b \leqq 2\sqrt{u}} b \psi\left(\frac{4u}{b}\right) u\, du = \sum_{b \leqq 2\sqrt{x}} b \int_{\mathrm{Max}\left(\frac{b^2}{4};1\right)}^x \psi\left(\frac{4u}{b}\right) u\, du = B \sum_{b \leqq 2\sqrt{x}} b x b$$

$$= B x^{\frac{5}{2}} \quad (35).$$

$$\int_0^x T(4u)\, T(u)\, du = \int_1^x T(4u)\, T(u)\, du + B \quad (2)$$

$$\frac{x^3}{3} + 4 \int_1^x \sum_{a \leqq 2\sqrt{u}} \frac{1}{a} \psi'\left(\frac{4u}{a}\right) \sum_{b \leqq \sqrt{u}} \frac{1}{b} \psi\left(\frac{u}{b}\right) u^2 du + 4 \int_1^x \sum_{a \leqq 2\sqrt{u}} \frac{1}{a} \psi\left(\frac{4u}{a}\right) \sum_{b \leqq \sqrt{u}} b \psi\left(\frac{u}{b}\right) u\, du$$

$$+ \int_1^x \sum_{a \leqq 2\sqrt{u}} a \psi\left(\frac{4u}{a}\right) \sum_{b \leqq \sqrt{u}} \frac{1}{b} \psi'\left(\frac{u}{b}\right) u\, du + \int_1^x \sum_{a \leqq 2\sqrt{u}} a \psi\left(\frac{4u}{a}\right) \sum_{b \leqq \sqrt{u}} b \psi\left(\frac{u}{b}\right) du + B x^{\frac{5}{2}} \log x$$

(110 bis 114).

Zieht man hierin die Summen aus den Integralen heraus und beachtet (108), so folgt (109).

Hilfssatz 17:

$$(115) \qquad \sum_{\substack{a,b=1 \\ (a,b)=1,\ a \equiv 1 \pmod 2}}^{\infty} \frac{1}{a^2 b^2} = 2.$$

Beweis:

$$(116) \qquad \sum_{\substack{a,b=1 \\ (a,b)=1,\ a \equiv 0 \pmod 2}}^{\infty} \frac{1}{a^2 b^2} = \sum_{\substack{a,b=1 \\ (a,b)=1 \\ a \equiv 0,\ b \equiv 1 \pmod 2}}^{\infty} \frac{1}{a^2 b^2} = \frac{1}{4} \sum_{\substack{a,b=1 \\ (a,b)=1,\ b \equiv 1 \pmod 2}}^{\infty} \frac{1}{a^2 b^2}$$

$$= \frac{1}{4} \sum_{\substack{a,b=1 \\ (a,b)=1,\ a \equiv 1 \pmod 2}}^{\infty} \frac{1}{a^2 b^2},$$

$$\frac{5}{2} = \sum_{\substack{a,b=1\\(a,b)=1}}^{\infty} \frac{1}{a^2 b^2} = \sum_{\substack{a,b=1\\(a,b)=1,\ a\equiv 1 \pmod 2}}^{\infty} \frac{1}{a^2 b^2} + \sum_{\substack{a,b=1\\(a,b)=1,\ a\equiv 0 \pmod 2}}^{\infty} \frac{1}{a^2 b^2}$$

$$= \frac{5}{4} \sum_{\substack{a,b=1\\(a,b)=1,\ a\equiv 1 \pmod 2}}^{\infty} \frac{1}{a^2 b^2} \qquad (39, 116).$$

Hilfssatz 18:

(117) $$J_{\mathrm{III}} = \frac{2}{3\pi^2} \sum_{\substack{a\leqq 2\sqrt{x}\\ b\leqq \sqrt{x}}} \frac{x^3 - r^3}{ab} \sum_{\substack{m,n=1\\ 4mb=na}}^{\infty} \frac{1}{mn}$$

(118) $$= \frac{11\pi^2}{432} x^3 + B x^{\frac{5}{2}} \log x.$$

Beweis:

$$J_{\mathrm{III}} = \frac{1}{9} \sum_{\substack{a\leqq 2\sqrt{x}\\ b\leqq \sqrt{x}}} \frac{x^3 - r^3}{ab} \frac{(a,4b)^2}{4ab} \qquad (117, 38)$$

(119) $$= \frac{1}{36} x^3 \sum_{\substack{a\leqq 2\sqrt{x}\\ b\leqq \sqrt{x}}} \frac{(a,4b)^2}{a^2 b^2} - \frac{1}{36} \sum_{\substack{a\leqq 2\sqrt{x}\\ b\leqq \sqrt{x}}} \frac{r^3 (a,4b)^2}{a^2 b^2}.$$

(120) $$\sum_{\substack{a\leqq 2\sqrt{x}\\ b\leqq \sqrt{x}}} \frac{(a,4b)^2}{a^2 b^2} = \sum_{d\leqq 2\sqrt{x}} d^2 \sum_{\substack{a\leqq 2\sqrt{x}\\ b\leqq \sqrt{x}\\ (a,4b)=d}} \frac{1}{a^2 b^2}.$$

$$\sum_{d\leqq 2\sqrt{x}} d^2 \sum_{a\leqq 2\sqrt{x}} \frac{1}{a^2} \sum_{\substack{b>\sqrt{x}\\ (a,4b)=d}} \frac{1}{b^2} = B \sum_{d\leqq 2\sqrt{x}} d^2 \sum_{a=1}^{\infty} \frac{1}{d^2 a^2} \sum_{b>\frac{\sqrt{x}}{d}} \frac{1}{d^2 b^2}$$

(121) $$= \frac{B}{\sqrt{x}} \sum_{d\leqq 2\sqrt{x}} \frac{1}{d} = B \frac{\log x}{\sqrt{x}}.$$

$$\sum_{d\leqq 2\sqrt{x}} d^2 \sum_{a>2\sqrt{x}} \frac{1}{a^2} \sum_{\substack{b=1\\ (a,4b)=d}}^{\infty} \frac{1}{b^2} = B \sum_{d\leqq 2\sqrt{x}} d^2 \sum_{a>\frac{2\sqrt{x}}{d}} \frac{1}{d^2 a^2} \sum_{b=1}^{\infty} \frac{1}{d^2 b^2}$$

(122) $$= \frac{B}{\sqrt{x}} \sum_{d\leqq 2\sqrt{x}} \frac{1}{d} = B \frac{\log x}{\sqrt{x}}.$$

(123) $$\sum_{d>2\sqrt{x}} d^2 \sum_{\substack{a,b=1\\ (a,4b)=d}}^{\infty} \frac{1}{a^2 b^2} = B \sum_{d>2\sqrt{x}} d^2 \sum_{a,b=1}^{\infty} \frac{1}{d^4 a^2 b^2} = B \sum_{d>2\sqrt{x}} \frac{1}{d^2} = \frac{B}{\sqrt{x}}.$$

$$\sum_{\substack{a\leqq 2\sqrt{x}\\ b\leqq\sqrt{x}}}\frac{(a,4b)^2}{a^2b^2}=\sum_{d=1}^{\infty}d^2\sum_{\substack{a,b=1\\(a,4b)=d}}^{\infty}\frac{1}{a^2b^2}+\frac{B\log x}{\sqrt{x}}\qquad(120\text{ bis }123)$$

$$=\sum_{d=1}^{\infty}d^2\left(\sum_{\substack{a,b=1\\(a,4b)=d\\a\equiv 1\,(\mathrm{mod}\,2)}}^{\infty}+\sum_{\substack{a,b=1\\(a,4b)=d\\a\equiv 2\,(\mathrm{mod}\,4)}}^{\infty}+\sum_{\substack{a,b=1\\(a,4b)=d\\a\equiv 0\,(\mathrm{mod}\,4)}}^{\infty}\right)\frac{1}{a^2b^2}+\frac{B\log x}{\sqrt{x}}$$

(124)
$$=U_1+U_2+U_3+\frac{B\log x}{\sqrt{x}}.$$

$$U_1=\sum_{d=1}^{\infty}d^2\sum_{\substack{a,b=1\\(a,b)=d\\a\equiv 1\,(\mathrm{mod}\,2)}}^{\infty}\frac{1}{a^2b^2}=\sum_{\substack{d=1\\d\equiv 1\,(\mathrm{mod}\,2)}}^{\infty}\frac{1}{d^2}\sum_{\substack{a,b=1\\(a,b)=1,\ a\equiv 1\,(\mathrm{mod}\,2)}}^{\infty}\frac{1}{a^2b^2}$$

(125)
$$=\frac{\pi^2}{8}\cdot 2=\frac{\pi^2}{4}\qquad(124,115).$$

In U_2 ist $(a,2b)=d$, $a\equiv d\equiv 2\,(\mathrm{mod}\,4)$, also $\left(\frac{a}{2},b\right)=\frac{d}{2}$, $\frac{a}{2}\equiv\frac{d}{2}\equiv 1\,(\mathrm{mod}\,2)$. Schreibe ich jetzt a,b,d für $\frac{a}{2},b,\frac{d}{2}$, so folgt

(126)
$$U_2=\sum_{\substack{d=1\\d\equiv 1\,(\mathrm{mod}\,2)}}^{\infty}4d^2\sum_{\substack{a,b=1\\(a,b)=d\\a\equiv 1\,(\mathrm{mod}\,2)}}^{\infty}\frac{1}{4a^2b^2}=\frac{\pi^2}{4}\qquad(124,125).$$

In U_3 ist $(a,4b)=d$, $a\equiv d\equiv 0\,(\mathrm{mod}\,4)$, also $\left(\frac{a}{4},b\right)=\frac{d}{4}$. Schreibe ich jetzt a,b,d für $\frac{a}{4},b,\frac{d}{4}$, so folgt

$$U_3=\sum_{d=1}^{\infty}16d^2\sum_{\substack{a,b=1\\(a,b)=d}}^{\infty}\frac{1}{16a^2b^2}=\sum_{d=1}^{\infty}\frac{1}{d^2}\sum_{\substack{a,b=1\\(a,b)=1}}^{\infty}\frac{1}{a^2b^2}$$

(127)
$$=\frac{\pi^2}{6}\,\frac{5}{2}=\frac{5\pi^2}{12}\qquad(124,39).$$

(128)
$$\sum_{\substack{a\leqq 2\sqrt{x}\\ b\leqq\sqrt{x}}}\frac{(a,4b)^2}{a^2b^2}=\frac{11\pi^2}{12}+\frac{B\log x}{\sqrt{x}}\qquad(124\text{ bis }127).$$

(108)
$$\sum_{\substack{a\leqq 2\sqrt{x}\\ b\leqq\sqrt{x}}}\frac{r^3(a,4b)^2}{a^2b^2}=B\sum_{a,b\leqq 2\sqrt{x}}\frac{(a^6+b^6)(a,b)^2}{a^2b^2}=B\sum_{a,b\leqq 2\sqrt{x}}\frac{a^6(a,b)^2}{a^2b^2}$$

$$=B\sum_{a,b\leqq 2\sqrt{x}}\frac{a^4(a,b)^2}{b^2}$$

$$=B\sum_{d\leqq 2\sqrt{x}}d^2\sum_{a,b\leqq\frac{2\sqrt{x}}{d}}\frac{d^4a^4}{d^2b^2}=B\sum_{d\leqq 2\sqrt{x}}d^4\sum_{a\leqq\frac{2\sqrt{x}}{d}}a^4=Bx^{\frac{5}{2}}\sum_{d\leqq 2\sqrt{x}}\frac{1}{d}$$

(129)
$$=Bx^{\frac{5}{2}}\log x.$$

$$J_{\mathrm{III}}=\frac{x^3}{36}\,\frac{11\pi^2}{12}+Bx^{\frac{5}{2}}\log x=\frac{11\pi^2}{432}x^3+Bx^{\frac{5}{2}}\log x\qquad(119,128,129).$$

Hilfssatz 19:

$$(130)\qquad \sum_{\substack{a\leqq 2\sqrt{x}\\ b\leqq \sqrt{x}}} \frac{1}{ab} \sum_{m,n=1}^{\infty} \frac{1}{mn} \left| \int_T^x \cos 2\pi u \left(\frac{4m}{a} + \frac{n}{b}\right) u^2\, du \right| = B x^{\frac{5}{2}},$$

$$(131)\qquad \sum_{\substack{a\leqq 2\sqrt{x}\\ b\leqq \sqrt{x}}} \frac{b}{a} \sum_{m,n=1}^{\infty} \frac{1}{mn} \left| \int_T^x \cos 2\pi u \left(\frac{4m}{a} + \frac{n}{b}\right) u\, du \right| = B x^{\frac{5}{2}},$$

$$(132)\qquad \sum_{\substack{a\leqq 2\sqrt{x}\\ b\leqq \sqrt{x}}} \frac{a}{b} \sum_{m,n=1}^{\infty} \frac{1}{mn} \left| \int_T^x \cos 2\pi u \left(\frac{4m}{a} + \frac{n}{b}\right) u\, du \right| = B x^{\frac{5}{2}},$$

$$(133)\qquad \sum_{\substack{a\leqq 2\sqrt{x}\\ b\leqq \sqrt{x}}} ab \sum_{m,n=1}^{\infty} \frac{1}{mn} \left| \int_T^x \cos 2\pi u \left(\frac{4m}{a} + \frac{n}{b}\right) du \right| = B x^{\frac{5}{2}}.$$

Beweis: Mit

$$U = x^2 \sum_{a,b\leqq 2\sqrt{x}} \frac{1}{ab} \sum_{m,n=1}^{\infty} \frac{1}{mn\left(\frac{m}{a}+\frac{n}{b}\right)} = B x^{\frac{5}{2}} \qquad (77, 78)$$

sind die linken Seiten von (130) bis (133)

$$= BU = B x^{\frac{5}{2}},$$

$$= Bx \sum_{a,b\leqq 2\sqrt{x}} \frac{b^2}{ab} \sum_{m,n=1}^{\infty} \frac{1}{mn\left(\frac{m}{a}+\frac{n}{b}\right)} = BU = B x^{\frac{5}{2}},$$

$$= Bx \sum_{a,b\leqq 2\sqrt{x}} \frac{a^2}{ab} \sum_{m,n=1}^{\infty} \frac{1}{mn\left(\frac{m}{a}+\frac{n}{b}\right)} = BU = B x^{\frac{5}{2}},$$

$$= B \sum_{a,b\leqq 2\sqrt{x}} \frac{a^2 b^2}{ab} \sum_{m,n=1}^{\infty} \frac{1}{mn\left(\frac{m}{a}+\frac{n}{b}\right)} = BU = B x^{\frac{5}{2}}.$$

Hilfssatz 20:

$$(134)\qquad x^2 \sum_{\substack{a\leqq 2\sqrt{x}\\ b\leqq \sqrt{x}}} \frac{b}{a} \sum_{\substack{m,n=1\\ 4mb=na}}^{\infty} \frac{1}{mn} = B x^{\frac{5}{2}} \log x,$$

$$(135)\qquad x^2 \sum_{\substack{a\leqq 2\sqrt{x}\\ b\leqq \sqrt{x}}} \frac{a}{b} \sum_{\substack{m,n=1\\ 4mb=na}}^{\infty} \frac{1}{mn} = B x^{\frac{5}{2}} \log x,$$

$$(136)\qquad x \sum_{\substack{a\leqq 2\sqrt{x}\\ b\leqq \sqrt{x}}} ab \sum_{\substack{m,n=1\\ 4mb=na}}^{\infty} \frac{1}{mn} = B x^{\frac{5}{2}} \log x.$$

Beweis: Die linke Seite von (134) ist

$$(137)\qquad = Bx^2 \sum_{a,b \leqq 2\sqrt{x}} \frac{b}{a} \sum_{\substack{m,n=1 \\ mb=na}}^{\infty} \frac{1}{mn} = Bx^2 \sum_{a,b \leqq 2\sqrt{x}} \frac{(a,b)^2}{a^2} \qquad (38)$$

$$(138)\qquad = Bx^2 \sum_{d \leqq 2\sqrt{x}} d^2 \sum_{a,b \leqq \frac{2\sqrt{x}}{d}} \frac{1}{d^2 a^2} = Bx^{\frac{5}{2}} \sum_{d \leqq 2\sqrt{x}} \frac{1}{d} = Bx^{\frac{5}{2}} \log x;$$

die linke Seite von (135) ist

$$(139)\qquad = Bx^2 \sum_{a,b \leqq 2\sqrt{x}} \frac{a}{b} \sum_{\substack{m,n=1 \\ mb=na}}^{\infty} \frac{1}{mn} = Bx^2 \sum_{a,b \leqq 2\sqrt{x}} \frac{b}{a} \sum_{\substack{m,n=1 \\ mb=na}}^{\infty} \frac{1}{mn} = Bx^{\frac{5}{2}} \log x \qquad (137, 138)$$

(ich habe a mit b, m mit n vertauscht); die linke Seite von (136) ist

$$= Bx \sum_{a,b \leqq 2\sqrt{x}} ab \sum_{\substack{m,n=1 \\ mb=na}}^{\infty} \frac{1}{mn} = Bx^2 \sum_{a,b \leqq 2\sqrt{x}} \frac{a}{b} \sum_{\substack{m,n=1 \\ mb=na}}^{\infty} \frac{1}{mn} = Bx^{\frac{5}{2}} \log x \qquad (139).$$

Hilfssatz 21:

$$(140)\qquad \sum_{\substack{a \leqq 2\sqrt{x} \\ b \leqq \sqrt{x}}} \frac{1}{ab} \sum_{\substack{m,n=1 \\ 4mb \neq na}}^{\infty} \frac{1}{mn} \left| \int_T^x \cos 2\pi u \left(\frac{4m}{a} - \frac{n}{b}\right) u^2 \, du \right| = Bx^{\frac{5}{2}} \log x,$$

$$(141)\qquad \sum_{\substack{a \leqq 2\sqrt{x} \\ b \leqq \sqrt{x}}} \frac{b}{a} \sum_{\substack{m,n=1 \\ 4mb \neq na}}^{\infty} \frac{1}{mn} \left| \int_T^x \cos 2\pi u \left(\frac{4m}{a} - \frac{n}{b}\right) u \, du \right| = Bx^{\frac{5}{2}} \log x,$$

$$(142)\qquad \sum_{\substack{a \leqq 2\sqrt{x} \\ b \leqq \sqrt{x}}} \frac{a}{b} \sum_{\substack{m,n=1 \\ 4mb \neq na}}^{\infty} \frac{1}{mn} \left| \int_T^x \cos 2\pi u \left(\frac{4m}{a} - \frac{n}{b}\right) u \, du \right| = Bx^{\frac{5}{2}} \log x,$$

$$(143)\qquad \sum_{\substack{a \leqq 2\sqrt{x} \\ b \leqq \sqrt{x}}} ab \sum_{\substack{m,n=1 \\ 4mb \neq na}}^{\infty} \frac{1}{mn} \left| \int_T^x \cos 2\pi u \left(\frac{4m}{a} - \frac{n}{b}\right) du \right| = Bx^{\frac{5}{2}} \log x.$$

Beweis: Mit

$$(144)\qquad U' = x^2 \sum_{a,b \leqq 4\sqrt{x}} \frac{1}{ab} \sum_{\substack{m,n=1 \\ mb \neq na}}^{\infty} \frac{1}{mn \left| \frac{m}{a} - \frac{n}{b} \right|} = Bx^{\frac{5}{2}} \log x \qquad (72, 73)$$

ist die linke Seite von (140)

$$= Bx^2 \sum_{\substack{a \leqq 2\sqrt{x} \\ b \leqq \sqrt{x}}} \frac{1}{ab} \sum_{\substack{m,n=1 \\ 4mb \neq na}}^{\infty} \frac{1}{mn \left| \frac{m}{a} - \frac{n}{4b} \right|} = BU' = Bx^{\frac{5}{2}} \log x \qquad (144)$$

(ich habe b durch $b'=4b$ ersetzt, auf $b'\equiv 0 \pmod 4$ verzichtet und dann wieder b für b' geschrieben — ebenso später); die linke Seite von (141)

$$= Bx \sum_{a,b \leqq 4\sqrt{x}} \frac{b}{a} \sum_{\substack{m,n=1 \\ mb \neq na}}^{\infty} \frac{1}{mn\left|\frac{m}{a}-\frac{n}{b}\right|} = BU' = Bx^{\frac{5}{2}} \log x \qquad (144);$$

die linke Seite von (142)

$$= Bx \sum_{a,b \leqq 4\sqrt{x}} \frac{a}{b} \sum_{\substack{m,n=1 \\ mb \neq na}}^{\infty} \frac{1}{mn\left|\frac{m}{a}-\frac{n}{b}\right|} = BU' = Bx^{\frac{5}{2}} \log x \qquad (144);$$

die linke Seite von (143)

$$= B \sum_{a,b \leqq 4\sqrt{x}} ab \sum_{\substack{m,n=1 \\ mb \neq na}}^{\infty} \frac{1}{mn\left|\frac{m}{a}-\frac{n}{b}\right|} = BU' = Bx^{\frac{5}{2}} \log x \qquad (144).$$

Hilfssatz 22:

$$(145) \qquad 4 \sum_{\substack{a \leqq 2\sqrt{x} \\ b \leqq \sqrt{x}}} \frac{1}{ab} \int_r^x \psi\left(\frac{4u}{a}\right) \psi\left(\frac{u}{b}\right) u^2\, du = \frac{11\pi^2}{432} x^3 + Bx^{\frac{5}{2}} \log x.$$

Beweis: Mit $\alpha = r$, $\beta = x$, $g(u) = u\psi\left(\frac{4u}{a}\right)$, $h(u) = u\psi\left(\frac{u}{b}\right)$,

$$(146) \qquad s_k(u) = -\frac{u}{\pi} \sum_{m=1}^{k} \frac{\sin 2\pi \frac{4m}{a} u}{m}, \qquad t_k(u) = -\frac{u}{\pi} \sum_{n=1}^{k} \frac{\sin 2\pi \frac{n}{b} u}{n}$$

sind die Voraussetzungen von Hilfssatz 10 erfüllt [3]). Daher liefern (83), (146), (117), (118), (140) und (130)

$$4 \sum_{\substack{a \leqq 2\sqrt{x} \\ b \leqq \sqrt{x}}} \frac{1}{ab} \int_r^x \psi\left(\frac{4u}{a}\right) \psi\left(\frac{u}{b}\right) u^2\, du$$

$$= \lim_{k=\infty} \frac{4}{\pi^2} \sum_{\substack{a \leqq 2\sqrt{x} \\ b \leqq \sqrt{x}}} \frac{1}{ab} \sum_{m,n=1}^{k} \frac{1}{mn} \int_r^x u^2 \sin 2\pi \frac{4m}{a} u \sin 2\pi \frac{n}{b} u\, du$$

$$= \lim_{k=\infty} \frac{2}{\pi^2} \sum_{\substack{a \leqq 2\sqrt{x} \\ b \leqq \sqrt{x}}} \frac{1}{ab} \Bigg(\sum_{\substack{m,n=1 \\ 4mb=na}}^{k} \frac{1}{mn} \int_r^x u^2\, du + \sum_{\substack{m,n=1 \\ 4mb \neq na}}^{k} \frac{1}{mn} \int_r^x \cos 2\pi u \left(\frac{4m}{a} - \frac{n}{b}\right) u^2\, du$$

$$- \sum_{m,n=1}^{k} \frac{1}{mn} \int_r^x \cos 2\pi u \left(\frac{4m}{a} + \frac{n}{b}\right) u^2\, du \Bigg) = \frac{11\pi^2}{432} x^3 + Bx^{\frac{5}{2}} \log x.$$

Hilfssatz 23:

$$\sum_{\substack{a\leqq 2\sqrt{x}\\ b\leqq\sqrt{x}}}\frac{b}{a}\int_r^x \psi\left(\frac{4u}{a}\right)\psi\left(\frac{u}{b}\right)u\,du = B\,x^{\frac{3}{2}}\log x. \tag{147}$$

Beweis: Mit $\alpha = r$, $\beta = x$, $g(u) = \sqrt{u}\,\psi\left(\frac{4u}{a}\right)$, $h(u) = \sqrt{u}\,\psi\left(\frac{u}{b}\right)$,

$$s_k(u) = -\frac{\sqrt{u}}{\pi}\sum_{m=1}^{k}\frac{\sin 2\pi\frac{4m}{a}u}{m}; \qquad t_k(u) = -\frac{\sqrt{u}}{\pi}\sum_{n=1}^{k}\frac{\sin 2\pi\frac{n}{b}u}{n} \tag{148}$$

sind die Voraussetzungen von Hilfssatz 10 erfüllt [3]). Daher liefern (83), (148), (134), (141) und (131)

$$\sum_{\substack{a\leqq 2\sqrt{x}\\ b\leqq\sqrt{x}}}\frac{b}{a}\int_r^x \psi\left(\frac{4u}{a}\right)\psi\left(\frac{u}{b}\right)u\,du$$

$$= \lim_{k=\infty}\frac{1}{\pi^2}\sum_{\substack{a\leqq 2\sqrt{x}\\ b\leqq\sqrt{x}}}\frac{b}{a}\sum_{m,n=1}^{k}\frac{1}{mn}\int_r^x u\sin 2\pi\frac{4m}{a}u\sin 2\pi\frac{n}{b}u\,du$$

$$= \lim_{k=\infty}\frac{1}{2\pi^2}\sum_{\substack{a\leqq 2\sqrt{x}\\ b\leqq\sqrt{x}}}\frac{b}{a}\Bigg(\sum_{\substack{m,n=1\\ 4mb=na}}^{k}\frac{1}{mn}\int_r^x u\,du + \sum_{\substack{m,n=1\\ 4mb\neq na}}^{k}\frac{1}{mn}\int_r^x \cos 2\pi u\left(\frac{4m}{a}-\frac{n}{b}\right)u\,du$$

$$-\sum_{m,n=1}^{k}\frac{1}{mn}\int_r^x \cos 2\pi u\left(\frac{4m}{a}+\frac{n}{b}\right)u\,du\Bigg) = B\,x^{\frac{3}{2}}\log x.$$

Hilfssatz 24:

$$\sum_{\substack{a\leqq 2\sqrt{x}\\ b\leqq\sqrt{x}}}\frac{a}{b}\int_r^x \psi\left(\frac{4u}{a}\right)\psi\left(\frac{u}{b}\right)u\,du = B\,x^{\frac{5}{2}}\log x. \tag{149}$$

Beweis: Genau wie der von Hilfssatz 23, unter Benutzung von (83), (148), (135), (142) und (132).

Hilfssatz 25:

$$\sum_{\substack{a\leqq 2\sqrt{x}\\ b\leqq\sqrt{x}}}ab\int_r^x \psi\left(\frac{4u}{a}\right)\psi\left(\frac{u}{b}\right)du = B\,x^{\frac{3}{2}}\log x. \tag{150}$$

Beweis: Mit $\alpha = r$, $\beta = x$, $g(u) = \psi\left(\frac{4u}{a}\right)$, $h(u) = \psi\left(\frac{u}{b}\right)$,

$$s_k(u) = -\frac{1}{\pi}\sum_{m=1}^{k}\frac{\sin 2\pi\frac{4m}{a}u}{m}, \qquad t_k(u) = -\frac{1}{\pi}\sum_{n=1}^{k}\frac{\sin 2\pi\frac{n}{b}u}{n} \tag{151}$$

sind die Voraussetzungen von Hilfssatz 10 erfüllt [3]). Daher liefern (83), (151), (136), (143) und (133)

$$\sum_{\substack{a\leqq 2\sqrt{x}\\ b\leqq\sqrt{x}}} ab\int_r^x \psi\left(\frac{4u}{a}\right)\psi\left(\frac{u}{b}\right)du$$

$$=\lim_{k=\infty}\frac{1}{\pi^2}\sum_{\substack{a\leqq 2\sqrt{x}\\ b\leqq\sqrt{x}}} ab\sum_{m,n=1}^{k}\frac{1}{mn}\int_r^x \sin 2\pi\frac{4m}{a}u\sin 2\pi\frac{n}{b}u\,du$$

$$=\lim_{k=\infty}\frac{1}{2\pi^2}\sum_{\substack{a\leqq 2\sqrt{x}\\ b\leqq\sqrt{x}}} ab\Bigg(\sum_{\substack{m,n=1\\ 4mb=na}}^{k}\frac{1}{mn}\int_r^x du+\sum_{\substack{m,n=1\\ 4mb\neq na}}^{k}\frac{1}{mn}\int_r^x \cos 2\pi u\left(\frac{4m}{a}-\frac{n}{b}\right)du$$

$$-\sum_{m,n=1}^{k}\frac{1}{mn}\int_r^x \cos 2\pi u\left(\frac{4m}{a}+\frac{n}{b}\right)du\Bigg)=Bx^{\frac{3}{2}}\log x.$$

Hilfssatz 26:

$$\int_0^x T(4u)\,T(u)\,du=\frac{144+11\pi^2}{432}x^3+Bx^{\frac{5}{2}}\log x. \tag{152}$$

Beweis: (109), (145), (147), (149), (150).

Beweis von (III):

$$\int_0^x \mathsf{P}_1^2(u)\,du=\frac{288+40\pi^2}{27}x^3-2^{11}\frac{144+11\pi^2}{432}\frac{x^3}{2^6}+Bx^{\frac{5}{2}}\log x \qquad (104, 152)$$

$$=\frac{2}{3}\pi^2x^3+Bx^{\frac{5}{2}}\log x.$$

Radość, den 10. April 1931.

(Eingegangen am 27. April 1931.)

Über den Cauchyschen Integralsatz.

Nachtrag zu der Arbeit Band 32, S. 476—480.

Von

Lothar Heffter in Freiburg i. B.

In Nr. 4 der genannten Arbeit habe ich mit der von mir 1902 eingeführten Doppelsummenmethode einen Beweis des Cauchyschen Integralsatzes unter den von Lichtenstein (Sitzungsber. der Berl. Math. Ges. 1910, S. 84 ff.) zuerst angegebenen Voraussetzungen geführt. Da aber die Schlüsse auf S. 480, wie ich unmittelbar nach Erscheinen der Arbeit bemerkte, in ihrer Kürze nicht völlig durchsichtig sind und die ausführliche Darlegung gewisse Schwierigkeiten bereitet, so möchte ich nachstehend zeigen, wie man diese durch eine vorangehende Umformung der a. a. O. benutzten Doppelsumme (3) überwindet.

Die Voraussetzungen von Lichtenstein lauten:

1. In einem Gebiet G, dem das achsenparallele Rechteck $R \equiv [(a\alpha)(b\beta)]$ mit Einschluß seines Randes angehört, ist für jedes x, y

$$\lim_{\delta \to 0} U(x, y, \delta) \equiv \lim_{\delta \to 0} \left[\frac{g(x+\delta, y) - g(x, y)}{\delta} - \frac{f(x, y+\delta) - f(x, y)}{\delta} \right] = 0, \tag{1}$$

ohne daß in dem ganzen Gebiet G *gleichmäßige* Nullkonvergenz von U vorausgesetzt wird.

2. $f(x, y)$ und $g(x, y)$ sind in G zweidimensional *stetige* Funktionen des Variablenpaares x, y.

3. $U(x, y, \delta)$ ist für alle x, y in G und alle positiven δ beschränkt (oder „gleichmäßig endlich"), d. h. $|U(x, y, \delta)| < M$, M eine von x, y, δ unabhängige, endliche, positive Zahl.

Zum Beweis sei das Rechteck R wieder in kongruente Quadrate zerlegbar und ε eine beliebig kleine positive Zahl. R werde zunächst in $m_0 n_0$ so kleine Quadrate mit der Seitenlänge

$$\delta = \frac{b-a}{m_0} = \frac{\beta-\alpha}{n_0} \tag{2}$$

geteilt, daß innerhalb jedes Quadrates die Schwankung von $f(x, y)$ und $g(x, y) < \varepsilon$ bleibt; das ist wegen der Stetigkeit, also gleichmäßigen Stetig-

keit von f und g in R stets möglich. Diese Eigenschaft bleibt natürlich beim Übergang zu noch kleineren Quadraten bestehen.

Da ferner aus der Stetigkeit von f und g auch die gleichmäßige Stetigkeit von $U(x, y, \delta)$ in R bei festem δ folgt, so kann eine positive ganze Zahl $\varkappa$ ($\geqq 2$) so bestimmt werden, daß durch Teilung jeder Quadratseite in $\varkappa$ gleiche Teile ein engeres Netz von mn so kleinen Quadraten entsteht, daß in jedem von ihnen die Schwankung von $U(x, y, \delta) < \frac{\varepsilon}{2}$ ist. Die Gitterpunkte dieses engeren Gitters seien x_μ, y_ν ($\mu = 0, 1, \ldots, m-1$; $\nu = 0, 1, \ldots, n-1$; $x_0 \equiv a$, $x_m \equiv b$, $y_0 \equiv \alpha$, $y_n \equiv \beta$), und die Quadratseite sei δ_1, so daß

$$(3) \qquad \delta_1 = \frac{\delta}{\varkappa} = \frac{b-a}{m} = \frac{\beta-\alpha}{n} = x_{\mu+1} - x_\mu = y_{\nu+1} - y_\nu.$$

Wie hier von m_0, n_0, δ zu m, n, δ_1, kann man von diesen unter Benutzung derselben Zahl ε zu m_1, n_1, δ_2 übergehen, usw. Die Zahlen $\delta, \delta_1, \delta_2, \ldots$ konvergieren also gegen Null, die Zahlenpaare m, n; m_1, n_1; ... gegen ∞.

Nun bilden wir die Doppelsumme

$$(4) \quad S_{mn} \equiv \sum_{\mu,\nu} [(f(x_\mu, y_\nu) - f(x_\mu, y_{\nu+\varkappa}))(x_{\mu+1} - x_\mu) + (g(x_{\mu+\varkappa}, y_\nu) - g(x_\mu, y_\nu))(y_{\nu+1} - y_\nu)]$$
$$(\mu = 0, 1, \ldots, m-\varkappa;\ \nu = 0, 1, \ldots, n-\varkappa),$$

so daß einerseits

$$(5) \qquad S_{mn} = \varkappa \sum_{\mu,\nu} U(x_\mu, y_\nu, \delta)(x_{\mu+1} - x_\mu)(y_{\nu+1} - y_\nu),$$

andererseits

$$(6) \quad S_{mn} = \sum_{\mu=0}^{m-\varkappa} [f(x_\mu, \alpha) + f(x_\mu, y_1) + \ldots + f(x_\mu, y_{\varkappa-1}) - f(x_\mu, y_{n-\varkappa+1}) - \ldots - f(x_\mu, \beta)](x_{\mu+1} - x_\mu)$$
$$- \sum_{\nu=0}^{n-\varkappa} [g(a, y_\nu) + g(x_1, y_\nu) + \ldots + g(x_{\varkappa-1}, y_\nu) - g(x_{m-\varkappa+1}, y_\nu) - \ldots - g(b, y_\nu)](y_{\nu+1} - y_\nu).$$

Fügt man hierzu:

$$(7) \quad A \equiv \sum_{\mu=0}^{m-\varkappa} [(\varkappa-1) f(x_\mu, \alpha) - f(x_\mu, y_1) - \ldots - f(x_\mu, y_{\varkappa-1}) + f(x_\mu, y_{n-\varkappa+1}) + \ldots + f(x_\mu, y_{n-1}) - (\varkappa-1) f(x_\mu, \beta)](x_{\mu+1} - x_\mu)$$
$$- \sum_{\nu=0}^{n-\varkappa} [(\varkappa-1) g(a, y_\nu) - g(x_1, y_\nu) - \ldots - g(x_{\varkappa-1}, y_\nu) + g(x_{m-\varkappa+1}, y_\nu) + \ldots + g(x_{m-1}, y_\nu) - (\varkappa-1) g(b, y_\nu)](y_{\nu+1} - y_\nu),$$

so ist

$$(8) \quad S_{mn} + A = \varkappa \sum_{\mu=0}^{m-\varkappa} [f(x_\mu, \alpha) - f(x_\mu, \beta)](x_{\mu+1} - x_\mu) - \varkappa \sum_{\nu=0}^{n-\varkappa} [g(a, y_\nu) - g(b, y_\nu)](y_{\nu+1} - y_\nu)$$

und wegen der Bestimmung von δ

$$(9) \qquad |A| < 2(\varkappa - 1)\varepsilon(b - a + \beta - \alpha) < 2\varkappa\varepsilon(b - a + \beta - \alpha),$$

$$(10) \qquad A = 2\vartheta\varkappa\varepsilon(b - a + \beta - \alpha), \qquad (-1 < \vartheta < +1).$$

Fügt man zu $S_{mn} + A$ weiter hinzu:

$$(11) \quad B \equiv \varkappa \sum_{\mu=m-\varkappa+1}^{m-1} [f(x, \alpha) - f(x_\mu, \beta)](x_{\mu+1} - x_\mu) - \varkappa \sum_{\nu=n-\varkappa+1}^{n-1} [g(a, y_\nu) - g(b, y_\nu)](y_{\nu+1} - y_\nu),$$

so ist

$$(12) \quad S_{mn} + A + B = \varkappa \sum_{\mu=0}^{m-1} [f(x_\mu, \alpha) - f(x_\mu, \beta)](x_{\mu+1} - x_\mu) - \varkappa \sum_{\nu=0}^{n-1} [g(a, y_\nu) - g(b, y_\nu)](y_{\nu+1} - y_\nu).$$

Dabei ist, wenn man (3) berücksichtigt und den größeren der Maximalwerte von $|f|$ und $|g|$ im Rechteck R mit N bezeichnet,

$$(13) \qquad |B| < \varkappa 4 N\delta,$$

$$(14) \qquad B = \vartheta' \varkappa 4 N\delta, \qquad (-1 < \vartheta' < +1).$$

Addiert man also zu (4) die Gleichungen (10) und (14) und ersetzt $S_{mn} + A + B$ durch den Ausdruck (12), so kann man durch $\varkappa$ kürzen und erhält die für den Beweis *entscheidende Gleichung*[1])

$$(15) \quad \sum_{\mu=0}^{m-1} [f(x_\mu, \alpha) - f(x_\mu, \beta)](x_{\mu+1} - x_\mu) - \sum_{\nu=0}^{n-1} [g(a, y_\nu) - g(b, y_\nu)](y_{\nu+1} - y_\nu) = \sum_{\mu,\nu} U(x_\mu, y_\nu, \delta)(x_{\mu+1} - x_\mu)(y_{\nu+1} - y_\nu) + 2\vartheta\varepsilon(b - a + \beta - \alpha) + 4\vartheta' N\delta.$$
$$(\mu = 0, 1, \ldots, m - \varkappa;\ \nu = 0, 1, \ldots, n - \varkappa).$$

Der Limes der *linken* Seite von (15) für $\delta_i \to 0$, m_i, $n_i \to \infty$ ist das in positivem Sinn über den Rand von R erstreckte Cauchysche Integral $\int(f\,dx + g\,dy)$. — *Rechts* konvergiert für $\delta \to 0$ das letzte Glied $\to 0$, das vorletzte ist mit ε beliebig klein, und wir brauchen nur noch zu zeigen, daß auch die Summe auf der rechten Seite von (15) für $\delta \to 0$ einen Limes hat, der mit ε beliebig klein ist.

Ist nun $\lambda_\varepsilon(m, n)$ die Anzahl *der* in dieser Summe auftretenden Gitterpunkte des engeren Gitters mit der Quadratseite δ_1, für die $|U(x_\mu, y_\nu, \delta)| \geqq \varepsilon$,

---

[1]) Die linke Seite dieser Gleichung ist identisch mit der a. a. O. S. 478 benutzten Doppelsumme $\sum R_{\mu\nu}$ und kann auch in der Form geschrieben werden

$$\sum_{\mu,\nu} U(x_\mu, y_\nu, \delta_1)(x_{\mu+1} - x_\mu)(y_{\nu+1} - y_\nu).$$

Die Gleichung stellt also eine Umformung der damaligen Ausgangssumme dar.

so ist nach Voraussetzung 3.:

$$(16)\quad \sum_{\mu,\nu} |U(x_\mu, y_\nu, \delta)|\,(x_{\mu+1} - x_\mu)(y_{\nu+1} - y_\nu) < (b-a)(\beta-\alpha)\left[\left(1 - \frac{\lambda_\varepsilon(m,n)}{mn}\right)\varepsilon + \frac{\lambda_\varepsilon(m,n)}{mn} M\right]$$
$$(\mu = 0, 1, \ldots, m-\varkappa;\ \nu = 0, 1, \ldots, n-\varkappa).$$

Es muß aber

$$(17)\quad \lim_{i\to\infty} \frac{\lambda_\varepsilon(m_i, n_i)}{m_i n_i} = 0$$

sein. Denn nach der Art der Gitterverengerung von Schritt zu Schritt ist in *allen* Punkten x, y der Quadrate mit der Seite δ_{i+1}, in deren linker unterer Ecke $|U(x_\mu, y_\nu, \delta_i)| \geqq \varepsilon$ ist, $|U(x, y, \delta_i)| > \frac{\varepsilon}{2}$. Der Gesamtinhalt dieser Quadrate beträgt

$$(18)\quad \frac{\lambda_\varepsilon(m_i, n_i)}{m_i n_i}(b-a)(\beta-\alpha).$$

Wäre also (17) nicht erfüllt, so könnte auch der Flächeninhalt (18) nicht gegen Null konvergieren, und dann griffe der Satz von Arzelà-Hartogs Platz (Math. Abhandl. H. A. Schwarz zu seinem 50jährigen Dr.-Jubiläum gewidmet, Berlin 1914, S. 57): „R sei ein Rechteck, $R_1, R_2, \ldots$ seien Teile von R, deren jeder aus einer endlichen Anzahl einander nicht überdeckender Teilrechtecke besteht mit Seiten, die denen von R parallel sind. Gibt es dann keinen Punkt, der unendlich vielen R_ν angehört, so konvergiert der Gesamtinhalt von R_ν mit wachsendem ν gegen Null." Es gäbe danach in unserm Rechteck R mindestens einen Punkt ξ, η, in dem für unendlich viele Werte δ_i, die gegen Null konvergierten, $|U(\xi, \eta, \delta_i)| > \frac{\varepsilon}{2}$ wäre — im Widerspruch mit der Voraussetzung 1.:

Damit ist aber gezeigt, daß der Limes von (16) für $\delta \to 0$ mit ε beliebig klein ist, und damit nach (15) der Cauchysche Integralsatz bewiesen.

Nach alledem, was in dieser und der vorangehenden Arbeit ausgeführt oder in Erinnerung gerufen wurde, scheint es wohl sicher, daß *jeder* mit Hilfe des Doppel*integrals* zu führende Beweis des Cauchyschen Integralsatzes auch mit der Doppel*summen*methode geführt werden kann, was ja auch sehr plausibel ist. — *Nicht* mit dem Doppelintegral bewiesen ist m. W. der Cauchysche Satz nur unter den Bedingungen von Goursat. Dieser Beweis dürfte in der Arbeit Gött. Nachr. 1903 seine einfachste Gestalt gewonnen haben.

Freiburg i. B., 2. Mai 1931.

(Eingegangen am 15. Mai 1931.)

Einige Sätze und Fragestellungen über Fourier-Koeffizienten.

Von

S. Sidon in Budapest.

Nach einem klassischen Riemann-Lebesgueschen Satze muß, damit die trigonometrische Reihe $\sum_{n=0}^{\infty}(a_n \cos nx + b_n \sin nx)$ eine Fourier-Reihe[1]) ist, $\lim a_n = \lim b_n = 0$ sein; diese Bedingung ist aber, wie bekannt, nicht auch hinreichend; a_n und b_n haben, um die Fourier-Konstanten einer integrierbaren Funktion zu sein, noch anderen, weniger expliziten Bedingungen zu genügen. Es ist aber vielleicht die Tatsache nicht ganz ohne Interesse, daß sich Indexfolgen $\lambda_1, \lambda_2, \ldots, \lambda_n, \ldots$ angeben lassen, für die die zugehörigen Fourier-Koeffizienten nur der einzigen Beschränkung $\lim a_{\lambda_n} = \lim b_{\lambda_n} = 0$ unterworfen sind. Es gilt nämlich der

Satz I. *Sind* $\varepsilon_1', \varepsilon_2', \ldots, \varepsilon_n', \ldots$ *und* $\varepsilon_1'', \ldots, \varepsilon_n'', \ldots$ *zwei beliebige reelle Nullfolgen und erfüllen die positiven ganzen Zahlen* $\lambda_1, \lambda_2, \ldots, \lambda_n, \ldots$ *für jedes n die Bedingung*

$$\frac{\lambda_{n+1}}{\lambda_n} > q > 1 \quad (q \text{ konst.}), \tag{1}$$

so gibt es eine Fourier-Reihe $\sum_{n=0}^{\infty}(a_n \cos nx + b_n \sin nx)$ *mit*

$$a_{\lambda_n} = \varepsilon_n', \qquad b_{\lambda_n} = \varepsilon_n''.$$

Beim Beweise dieses Satzes stützen wir uns auf

Satz II[2]). *Sind* $\varepsilon_1', \varepsilon_2', \ldots, \varepsilon_n', \ldots$ *und* $\varepsilon_1'', \varepsilon_2'', \ldots, \varepsilon_n''$ *zwei beliebige reelle beschränkte Folgen und erfüllt die Indexfolge* $\lambda_1, \lambda_2 \ldots, \lambda_n, \ldots$ *die*

[1]) Fourier-Reihe bedeutet hier immer die Fourier-Reihe einer im Lebesgueschen Sinne integrierbaren Funktion.

[2]) Für den speziellen Fall, daß $\lambda_n > 2\sum_{k=1}^{n-1}\lambda_k$ erfüllt ist, ist dieser Satz schon in der Arbeit des Herrn F. Riesz, Über die Fourier-Koeffizienten einer stetigen Funktion von beschränkter Schwankung [Math. Zeitschr. 2 (1918), S. 312–315], enthalten.

Bedingung (1), *so gibt es eine für* $r<1$ *konvergente positive harmonische Entwicklung* $\sum_{n=0}^{\infty}(a_n\cos nx+b_n\sin nx)r^n$ [3]) *mit*

$$a_{\lambda_n}=\varepsilon_n',\qquad b_{\lambda_n}=\varepsilon_n''.$$

Beweis. Ohne Beschränkung der Allgemeinheit kann $|\varepsilon_n'|<\frac{1}{2}$, $|\varepsilon_n''|<\frac{1}{2}$ für jedes n vorausgesetzt werden. Wird die positive ganze Zahl k derart bestimmt, daß die Ungleichungen

$$q^k>3,\qquad 1+\frac{1}{q^k-2}<q$$

bestehen, dann ist in dem nichtnegativen trigonometrischen Polynom

$$P_{mr}(x)=\prod_{i=0}^{m}(1+\varepsilon'_{ki+r}\cos\lambda_{ki+r}x+\varepsilon''_{ki+r}\sin\lambda_{ki+r}x)=\sum_{i=0}^{M_r}(a_{ir}\cos ix+b_{ir}\sin ix)$$
$$(0\leqq r\leqq k-1),$$

wo $M_r=(m+1)\left(\frac{km}{2}+r\right)$, $a_{\lambda_{kl+r}}=\varepsilon'_{kl+r}$, $b_{\lambda_{kl+r}}=\varepsilon''_{kl+r}$. In der Folge der Polynome $P_{1r}(x), P_{2r}(x),\ldots,P_{nr}(x),\ldots$ ist jedes Glied ein Abschnitt der folgenden [4]). Die harmonische Entwicklung $\sum_{r=0}^{k-1}\sum_{i=0}^{\infty}(a_{ir}\cos ix+b_{ir}\sin ix)\varrho^i$ hat die in zu beweisendem Satze behauptete Eigenschaft.

Der Beweis des Satzes I läßt sich nun so führen: Es seien $q_1', q_2',\ldots$ und $q_2'',\ldots,q_n'',\ldots$ zwei konvexe Nullfolgen, für welche $|\varepsilon_n'|<q_{\lambda_n'}$, $|\varepsilon_n''|<q''_{\lambda_n}$ gilt für jedes n.[5]) Nach Satz II existiert dann eine positive harmonische Entwicklung $\sum_{n=0}^{\infty}(c_n\cos nx+d_n\sin nx)r^n$, in welcher $c_{\lambda_n}=\frac{\varepsilon_n'}{q'_{\lambda_n}}$, $d_{\lambda_n}=\frac{\varepsilon_n''}{q''_{\lambda_n}}$ ist. $\sum_{n=0}^{\infty}(a_n\cos nx+b_n\sin nx)$, wo $a_n=c_nq_n'$, $b_n=d_nq_n''$, ist eine Fourier-Reihe von der gewünschten Beschaffenheit [6]).

---

[3]) Aus der Existenz einer solchen Entwicklung folgt offenbar auch die unendlich vieler.

[4]) Siehe außer der zitierten Arbeit des Herrn F. Riesz die Noten des Verfassers. I. Ein Satz über die absolute Konvergenz von Fourier-Reihen, in denen sehr viele Glieder fehlen; II. Verallgemeinerung eines Satzes über die absolute Konvergenz von Fourier-Reihen mit Lücken [Math. Annalen **96** (1926), S. 418–419, und **97** (1927), S. 675–676].

[5]) Die Existenz solcher Folgen folgt aus Hilfssatz II meiner Arbeit: Reihentheoretische Sätze und ihre Anwendungen in der Theorie der Fourier-Reihen, Math. Zeitschr. **10** (1921), S. 121–127.

[6]) Daß $\sum_{n=0}^{\infty}a_n\cos nx+b_n\sin nx)$ eine Fourier-Reihe ist, folgt aus dem Youngschen Satze: Die „Klasse" einer Fourier-Reihe $\sum_{n=0}^{\infty}(a_n\cos nx+b_n\sin nx)$ bleibt durch die Komposition mit der durch die gliedweise Differentiation der Fourier-Reihe einer

(Fortsetzung der Fußnote [6]) auf nächster Seite.)

Es ist mir bisher nicht gelungen, in den soeben bewiesenen Sätzen die Bedingung (1) durch eine wesentlich allgemeinere zu ersetzen.

Ich befasse mich hier noch mit den Fourier-Koeffizienten stetiger Funktionen.

Herr Gronwall bewies, ein Resultat des Herrn Carleman verschärfend [7]), daß, wenn $\varphi(x)$ eine beliebige reelle Funktion mit $\lim_{x=0} \varphi(x) = \infty$ ist, stetige Funktionen mit der Fourier-Reihe $\sum_{n=0}^{\infty}(a_n \cos nx + b_n \sin nx)$ existieren, für welche $\sum_{n=0}^{\infty}(a_n^2 + b_n^2)\,\varphi(a_n^2 + b_n^2)$ divergiert. Hiermit ist aber die Frage noch nicht entschieden: Sind die Fourier-Koeffizienten stetiger Funktionen nur der einzigen quantitativen Beschränkung unterworfen, daß ihre Quadratsumme zu konvergieren hat?; d. h., lassen sich zu den Folgen $\alpha_1, \alpha_2, \ldots, \alpha_n, \ldots$ und $\beta_1, \ldots, \beta_n, \ldots$, wenn $\sum_{n=0}^{\infty}(\alpha_n^2 + \beta_n^2)$ konvergiert, immer stetige Funktionen angeben, deren Fourier-Koeffizienten den Ungleichungen

$$|a_n| > |\alpha_n|, \quad |b_n| > \beta_n \qquad (n = 1, 2, \ldots)$$

genügen?

Äquivalent mit dieser Frage ist die folgende: Sind die Faktorfolgen $\gamma_1, \gamma_2, \ldots, \gamma_n, \ldots$ mit konvergenter Quadratsumme die einzigen, welche die Fourier-Reihe einer beliebigen stetigen Funktion in eine absolut konvergente trigonometrische Reihe $\sum_{n=1}^{\infty}\gamma_n(a_n \cos nx + b_n \sin nx)$ überführen? Oder: Muß, wenn sämtliche trigonometrische Reihen $\sum_{n=0}^{\infty}\varepsilon_n \gamma_n \cos nx$, wo

---

Funktion von beschränkter Schwankung entstandenen Reihe $\sum_{n=0}^{\infty}(\alpha_n \cos nx + \beta_n \sin nx)$ unverändert. Hierbei ist unter Komposition die Bildung der Reihe

$$\sum_{n=0}^{\infty}[(\alpha_n a_n + \beta_n b_n)\cos nx + (\alpha_n b_n - a_n \beta_n)\sin nx]$$

zu verstehen. [W. H. Young, On Fourier-series and fonctions of bounded variation, Lond. Royal Soc. Proc. 88 (1913), p. 561–568; M. Fekete, Über Faktorenfolgen, welche die Klasse einer Fourier-Reihe unverändert lassen, Acta litt. ac. scient. Univ. Hung. 1923, S. 148–166.] In unserem Falle handelt es sich um die Klasse L (Klasse der nach Lebesgue integrierbaren Funktionen) der Fourier-Reihen $\sum_{n=0}^{\infty} q_n' \cos nx$, $\sum_{n=0}^{\infty} q_n'' \cos nx$ [W. H. Young, The Fourier-series of bounded fonctions, Proc. Lond. Mat. Soc. 12 (1913), p. 44–45], die mit $\sum_{n=0}^{\infty} c_n \cos nx$, bzw. $\sum_{n=0}^{\infty} d_n \sin nx$ komponiert werden.

[7]) T. Carleman, Über die Fourier-Koeffizienten einer stetigen Funktion, Acta Math. 1 (1918), S. 377–384; T. H. Gronwall, On the Fourier-coefficients of a continous fonctions, Amer. M. S. Bull. 27 (1921), p. 320–321.

die ε_n voneinander unabhängig die Werte 1 und -1 annehmen können, Fourier-Reihen sind, $\sum_{n=0}^{\infty} \gamma_n^2$ konvergieren[8])?

Aus der bejahenden Beantwortung unserer Fragestellungen würde, wenn $\lambda_1, \lambda_2, \ldots, \lambda_n, \ldots$ wieder eine die Bedingung (1) erfüllende Indexfolge ist, folgen: 1. Durch Kombination mit Satz II dieser Note, daß es, wenn $\sum_{n=1}^{\infty} (\varepsilon_n'^2 + \varepsilon_n''^2)$ konvergiert, stetige Funktionen mit der Fourier-Reihe $\sum_{n=0}^{\infty} (a_n \cos nx + b_n \sin nx)$ gibt, für welche

$$a_{\lambda_n} = \varepsilon_n', \qquad b_{\lambda_n} = \varepsilon_n''$$

ist[9]). 2. Durch Kombination mit dem in meiner unter [4]) zitierten Note II. bewiesenen Satze, daß die trigonometrische Reihe $\sum_{n=0}^{\infty} a_n \cos \lambda_n x$ nur dann eine Fourier-Reihe ist, wenn $\sum_{n=0}^{\infty} a_n^2$ konvergiert. Läßt sich nur eines dieser Korollare widerlegen, so folgt daraus die verneinende Beantwortung unserer Fragestellungen.

Es sei hier noch erwähnt, daß es genügt, unsere Fragestellungen für die Klasse der beschränkten Funktionen zu erledigen; der Übergang zu den stetigen Funktionen läßt sich dann auf analoge Weise vollziehen, wie dies beim Beweise des Satzes I von den positiven trigonometrischen Entwicklungen zu den Fourier-Reihen geschehen ist.

[8]) M. Fekete, loc. cit., wo die Umkehrung des oben angewandten Youngschen Satzes bewiesen wird, nämlich, daß die Reihen $\sum_{n=0}^{\infty} (\alpha_n \cos nx + \beta_n \sin nx)$ von derselben Beschaffenheit die einzigen sind, mit welchen komponiert die Klasse einer Fourier-Reihe unverändert bleibt. Sämtliche $\sum_{n=0}^{\infty} \varepsilon_n \gamma_n \cos nx$ müssen also die gliedweisen Derivierten der Fourier-Reihen von Funktionen von beschränkter Schwankung sein, woraus dann weiter folgt, daß sie Fourier-Reihen sind.

[9]) Hierbei bedienen wir uns wieder des Youngschen Satzes, der jetzt auf die Klasse der stetigen Funktionen angewendet wird.

(Eingegangen am 1. August 1928.)

Ein Satz über Fouriersche Reihen mit Lücken.

Von

S. Sidon in Budapest.

Als Beantwortung eines Spezialfalles einer in einer früheren Arbeit[1]) gestellten Frage beweise ich jetzt den

Satz I. *Erfüllt die unendliche Folge positiver, ganzer Zahlen* $\lambda_1, \lambda_2, \ldots, \lambda_n \ldots$ *für jedes* n *die Bedingung*

$$\frac{\lambda_{n+1}}{\lambda_n} > q > 1 \qquad (q \text{ konstant}), \tag{1}$$

so ist, wenn die trigonometrische Reihe $\sum_{n=1}^{\infty}(a_n \cos \lambda_n x + b_n \sin \lambda_n x)$ *eine Fourier-Reihe ist,* $\sum_{n=1}^{\infty}(a_n^2 + b_n^2)$ *konvergent.*

Beim Beweise dieses Satzes stützen wir uns auf folgende Hilfssätze:

Hilfssatz I. Ist $\sum_{n=1}^{\infty}(a_n \cos \lambda_n x + b_n \sin \lambda_n x)$[2]) eine Fouriersche Reihe, so sind es auch alle aus ihr durch Weglassung von Gliedern entstandenen trigonometrischen Reihen $\sum(a_m \cos \lambda_m x + b_m \sin \lambda_m x)$. Es gilt dies nämlich von sämtlichen $\sum_{n=1}^{\infty}\varepsilon_n(a_n \cos \lambda_n x + b_n \sin \lambda_n x)$, wo die ε_n voneinander unabhängig die Werte 1 oder -1 annehmen können[3]).

---

[1]) Einige Sätze und Fragestellungen über Fouriersche Koeffizienten. Dort wird auch der hier gelöste Spezialfall des allgemeinen Problems besonders erwähnt.

[2]) Mit $\lambda_1, \lambda_2, \ldots, \lambda_n \ldots$ ist hier immer eine die Bedingung 1. erfüllende Folge von positiven ganzen Zahlen bezeichnet.

[3]) Siehe die Noten des Verfassers: I. Ein Satz über die absolute Konvergenz von Fourier-Reihen, in denen sehr viele Glieder fehlen; II. Verallgemeinerung eines Satzes über die absolute Konvergenz von Fourier-Reihen mit Lücken. Math. Annalen **96** (1926), S. 418—419, spez. S. 419, Satz II und **97** (1927), S. 675—676.

Hilfssatz II. Ist $\sum_{n=1}^{\infty}(a_n \cos \lambda_n x + b_n \sin \lambda_n x)$ eine Fouriersche Reihe, so läßt sich dies auch von ihrer Konjugierten behaupten[4]).

Hilfssatz III. Ist $q > 2$, so gilt für jedes komplexe Polynom $P(x) = \sum_{k=1}^{n} c_k e^{i\lambda_k x}$

$$\frac{\int_0^{2\pi} |P(x)|^4 dx}{\left(\int_0^{2\pi} |P(x)|^2 dx\right)^2} \leqq \frac{1}{\pi}.$$

Beweis. Es ist

$$P^2(x) = \sum_{k=1,\, l=1}^{n,\, n} c_k c_l e^{i(\lambda_k + \lambda_l)x}.$$

Eine Gleichung

$$\lambda_{k_1} + \lambda_{l_1} = \lambda_{k_2} + \lambda_{l_2}$$

kann nur bestehen, wenn $\lambda_{k_1} = \lambda_{k_2}$, $\lambda_{l_1} = \lambda_{l_2}$ oder $\lambda_{k_1} = \lambda_{l_2}$, $\lambda_{l_1} = \lambda_{k_2}$ ist. Also ist in der Tat

$$\int_0^{2\pi} |P(x)|^4 dx = 2\pi \sum_{k=1,\, l=1}^{n,\, n} |c_k c_l|^2 < 4\pi \left(\sum_{k=1}^{n} |c_k|^2\right)^2 = \frac{1}{\pi}\left(\int_0^{2\pi} |P(x)|^2 dx\right)^2.$$

Hilfssatz IV. Ist $f(x)$ eine im Intervalle $a \leqq x \leqq b$ stetige nichtnegative Funktion, $t_1 > 0$, $t_2 > 0$, $t_1 > t_2$, so gilt die Ungleichung

$$\frac{\int_a^b [f(x)]^{2t_1} dx}{\left[\int_a^b [f(x)]^{t_1} dx\right]^2} \geqq \frac{\int_a^b [f(x)]^{2t_2} dx}{\left[\int_a^b [f(x)]^{t_2} dx\right]^2}.$$

Beweis. Die Funktion $\varphi(t) = \dfrac{\sum_{i=1}^{n} x_i^{2t}}{\left(\sum_{i=1}^{n} x_i^t\right)^2}$, wo die x_i nichtnegative Konstanten bedeuten, ist im Intervalle $0 < t < \infty$ monoton wachsend, weil dort überall

$$\varphi'(t) = \frac{\sum_{i=1,\, k=1}^{n,\, n} (\log x_i - \log x_k) x_i^t x_k^t (x_i^t - x_k^t)}{\left(\sum_{i=1}^{n} x_i^t\right)^3} > 0$$

ist. Der zu beweisende Hilfssatz ergibt sich hieraus durch Grenzübergang.

---

[4]) Läßt sich auch nach der in loc. cit. [3]) angewandten Methode beweisen.

Beweis des Satzes I. Es werde, was nach Hilfssatz I keine Beschränkung der Allgemeinheit ist, $q > 2$ vorausgesetzt. Nach Hilfssatz IV gilt dann, wenn $t_1 = 2$, $t_2 = 1$ gesetzt wird,

$$\frac{\int_0^{2\pi} |S_{\lambda_n}(x)|^4 dx}{\left(\int_0^{2\pi} |S_{\lambda_n}(x)|^2 dx\right)^2} > \frac{\int_0^{2\pi} |S_{\lambda_n}(x)|^2 dx}{\left(\int_0^{2\pi} |S_{\lambda_n}(x)|\, dx\right)^2},$$

wo $S_k(x)$ das k-te Fejérsche Mittel von $\sum_{n=1}^{\infty}(a_n + i\,b_n)\,e^{i\lambda_n x}$ bedeutet. Aus Hilfssatz III folgt dann weiter

$$\frac{\int_0^{2\pi} |S_{\lambda_n}(x)|^2 dx}{\left(\int_0^{2\pi} |S_{\lambda_n}(x)|\, dx\right)^2} < \frac{1}{\pi}.$$

Bleibt also $\int_0^{2\pi} |S_{\lambda_n}(x)|\, dx$ unter einer von n unabhängigen Schranke, so läßt sich dies auch von $\int_0^{2\pi} |S_{\lambda_n}(x)|^2 dx$, also auch von $\sum_{k=1}^{n}(a_k^2 + b_k^2)$ behaupten, womit Satz I bewiesen ist[5]).

Ich erwähne hier noch ein auf die Fourierschen Konstanten stetiger Funktionen bezügliches Korollar des Satzes I.

Satz II. *Ist* $\sum_{n=1}^{\infty}{}' |\alpha_n(a_{\lambda_n} + b_{\lambda_n})|$, *wo die* α_n *Konstanten* a_n, b_n *die Fourierschen Koeffizienten einer beliebigen stetigen Funktion sind, immer konvergent, so muß* $\sum_{n=1}^{\infty}{}' |\alpha_n|^2$ *konvergieren*[6]).

(Eingegangen am 7. September 1929.)

Zusatz während der Korrektur.

Satz I wurde inzwischen auch von Herrn A. Zygmund gefunden (Sur les séries trigonométriques lacunaires. The Journal of the Lond. Math. Soc. 5, Part 2, April 1930, S. 138—145, received december 1929), der ihn

[5]) Es ergibt sich sogar, daß auch die vierte Potenz der durch

$$\sum_{n=1}^{\infty}(a_n \cos \lambda_n x + b_n \sin \lambda_n x)$$

definierten Funktion integrierbar ist.

[6]) Siehe loc. cit. [1]), besonders die Fußnoten [6]) und [8]).

durch Kombination zweier mit Hilfe der in meinen unter [3]) zitierten Noten angewandten Methode abgeleiteten — mit den obigen Hilfssätzen I und II identischen — Hilfssätze mit einem Littlewoodschen Satze erhält (J. E. Littlewood, On the mean values of power series, Proc. of the Lond. Math. Soc. Ser. 2, **25**, Part 5, 1926, S. 328—337).

Eine leichte Modifikation des obigen Beweises des Satzes I ergibt, daß sich in demselben die Bedingung 1 durch die allgemeinere ersetzen läßt: In der Folge $\lambda_{11} = 2\lambda_1$, $\lambda_{12} = \lambda_1 + \lambda_2$, $\lambda'_{12} = \lambda_1 - \lambda_2, \ldots, \lambda_{mn} = \lambda_m + \lambda_n$, $\lambda'_{mn} = \lambda_m - \lambda_n \ldots$ bleibt die Anzahl der gleichen Glieder unter einer von m und n unabhängigen Schranke. Diese Bedingung ist tatsächlich wesentlich allgemeiner als 1.; es lassen sich nämlich dieselbe erfüllende λ-Folgen mit $\lambda_n = O(n^4)$ angeben.

(Eingegangen am 15. September 1931.)

Ein Satz über die Fourierschen Reihen stetiger Funktionen.

Von

S. Sidon in Budapest.

In der vorliegenden Note zeige ich, wie sich aus einem wichtigen Satze des Herrn Littlewood [1]) sehr einfach die Beantwortung einer von mir in einer früheren Arbeit [2]) gestellten Frage ergibt. Ich beweise nämlich den

Satz I. *Die Faktorenfolgen* $\gamma_1, \gamma_2, \ldots, \gamma_n$ *mit konvergenter* $\sum_{n=1}^{\infty} |\gamma_n^2|$ *sind die einzigen, welche die Fouriersche Reihe einer beliebigen stetigen Funktion* $\sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx)$ *in eine absolut konvergente trigonometrische Reihe* $\sum_{n=1}^{\infty} \gamma_n (a_n \cos nx + b_n \sin nx)$ *überführen.*

Der Littlewoodsche Satz, auf dem der Beweis des Satzes I beruht, lautet:

Ist die Potenzreihe $\sum_{n=0}^{\infty} c_n z^n$ *für* $|z|<1$ *konvergent,* $\sum |c_n|^2$ *divergent, so läßt sich eine reelle Folge* $\alpha_1, \alpha_2, \ldots, \alpha_n, \ldots$ *derart angeben, daß* $\lim_{r=1} \int_0^{2\pi} \left| \sum_{n=0}^{\infty} c_n e^{i\alpha_n} (re^{i\varphi})^n \right| d\varphi = \infty$ *ist.*

Beweis des Satzes I. Ist $\gamma_1, \gamma_2, \ldots, \gamma_n \ldots$ eine (nicht notwendig reelle) Folge von der erwähnten Beschaffenheit, so ist es auch $\gamma_1 e^{i\alpha_1}, \gamma_2 e^{i\alpha_2}, \ldots, \gamma_n e^{i\alpha_n} \ldots$. Dann sind aber die trigonometrischen Reihen $\sum_{n=1}^{\infty} \gamma_n e^{i\alpha_n} \cos nx$ und $\sum_{n=1}^{\infty} \gamma_n e^{i\alpha_n} \sin nx$ durch gliedweise Differentiation der

---

[1]) J. E. Littlewood, On the mean value of power series, Proc. London Math. Soc. **25** (1926), (2), S. 328–337. Von dem Littlewoodschen Satze nahm ich durch die unter [6]) zitierte Arbeit des Herrn A. Zygmund Kenntnis.

[2]) Einige Sätze und Fragestellungen über Fourier-Koeffizienten.

Fourierschen Reihen von Funktionen von beschränkter Schwankung entstanden[3]). $\int_0^{2\pi} \left| \sum_{n=1}^{\infty} \gamma_n e^{i\alpha_n} (re^{i\varphi})^n \right| d\varphi$ bleibt also für $r<1$ beschränkt, woraus nach dem Littlewoodschen Satze die Konvergenz von $\sum_{n=1}^{\infty} |\gamma_n|^2$ folgt, q. e. d.

Aus Satz I ergibt sich auf eine geläufige Weise der bekannte Satz des Herrn Carleman über den Konvergenzexponenten der Fourierschen Konstanten stetiger Funktionen[4]).

Aus Satz I folgt auch die folgende Verschärfung des Littlewoodschen Satzes[5]).

Satz II. *Sind sämtliche trigonometrische Reihen*

$$\sum_{n=1}^{\infty} \varepsilon_n (a_n \cos nx + b_n \sin nx),$$

wo die ε_n voneinander unabhängig die Werte 1 und -1 annehmen können, Fouriersche Reihen, so muß $\sum_{n=1}^{\infty} (a_n^2 + b_n^2)$ konvergieren.

Ein Korollar von Satz II ist[6]):

Ist die trigonometrische Reihe $\sum_{k=1}^{\infty} (a_k \cos n_k x + b_k \sin n_k x)$, wo von einem gewissen k an die Bedingung $\frac{n_{k+1}}{n_k} > q > 1$ erfüllt ist, eine Fouriersche Reihe, so muß $\sum (a_k^2 + b_k^2)$ konvergieren.

---

[3]) loc. cit.[2]) Fußnote [6]) und [8]).

[4]) Siehe das Zitat in meiner oben angeführten Arbeit. Auch die dort erwähnte, von Gronvall herrührende sowie auch die von mir gegebene Verschärfung des Carlemanschen Satzes (Ein Satz über trigonometrische Polynome mit Lücken, erscheint im Journal für die reine und angewandte Mathematik) ergibt sich aus Satz I. Die letztere lautet: Zu einer beliebigen Indexfolge $n_1, n_2, \ldots, n_k, \ldots$ läßt sich eine stetige Funktion angeben, bei welcher $\sum_{k=1}^{\infty} (|a_{n_k}|^{2-\delta} + |b_{n_k}|^{2-\delta})$ divergiert, wo $\delta > 0$ beliebig ist.

[5]) loc. cit.[3]).

[6]) Vgl. loc. cit.[2]). Diesen Satz habe ich elementar, ohne Anwendung des Littlewoodschen Satzes schon in meiner früheren Arbeit, „Ein Satz über Fourierreihen mit Lücken", bewiesen (eingesandt der Math. Zeitschr. am 5. September 1929). Er wurde inzwischen auch von Herrn A. Zygmund gefunden. (Sur les séries trigonométriques lacunaires. The Journal of the London Math. Soc. 5, Part 2, No. 18, pp. 138—145. Received 31. Dec. 1929, read 12. Dec. 1929 (sic).) Siehe auch den meiner soeben zitierten Arbeit hinzugefügten Zusatz.

(Eingegangen am 5. Juli 1930.)

Orthogonale Systeme elliptischer Funktionen.

Von

Wilhelm Maier in Frankfurt a. M.

Die Aufspaltung analytischer Funktionen in konvergente Reihen, wie sie von der Praxis gefordert und von der Theorie der partiellen Differentialgleichungen in einheitlicher Weise behandelt wurde, führte zum Begriff der Eigenfunktion linearer Differentialgleichungen. Neben die Theorie der Potenzreihen und trigonometrischen Reihen traten damit Entwickelungen nach Legendreschen Polynomen, nach Zylinderfunktionen und anderen Elementen, die der Laplaceschen Differentialgleichung $\Delta u = 0$ genügen. Wir knüpfen an die sogenannten Laméschen Funktionen an, die ersonnen wurden, um die Temperaturverteilung im Ellipsoid zu beschreiben. Daß deren Orthogonalitätseigenschaft zunächst übersehen werden konnte, erscheint heute verständlich aus zwei Gründen. Einmal setzten die Methoden Riemanns in der Funktionentheorie sich nur langsam durch; sodann fanden die Bemühungen Hermites und Kroneckers, welche als Grundlage aller periodischen Erscheinungen die „elliptischen Funktionen zweiter Art" benutzten, nicht genügend Nachfolge.

Wir betrachten Kroneckers s-Funktion als das theoretisch einfachste Gebilde im Rahmen der elliptischen Funktionen[1]). Sie ist mit s-Funktionen rational abgeänderter Argumente linear verknüpft (§ 1), führt auf finitem Wege zur $\wp$- und ϑ-Funktion (§§ 2—4) und genügt einem quadratischen Additionssatz (§ 5), dessen differentieller Grenzfall die gewünschte Orthogonalität belasteter s-Funktionen in Evidenz setzt (§ 6).

Die gefundenen Eigenfunktionen sind ϑ-Quotienten, die zugehörigen Eigenwerte Nullstellen von $\vartheta'\left(\frac{u}{\omega_1}\right) = 0$. Als einfaches Beispiel wird ab-

[1]) Bernoullische Polynome und elliptische Funktionen, Crelles Journal **164** (1931): Zur Theorie der elliptischen Funktionen, Theorie der t-Funktionen, Math. Annalen **104** (1931); Elementary properties of the t_ν-functions, Transact. Amer. Math. Soc. **32** (1930).

schließend eine Konstante nach ϑ-Quotienten entwickelt (§ 7); mit geeigneten „Amplituden“ c_n und „Phasen“ v_n gilt z. B.

$$\vartheta(v) = \sum_{-\infty}^{\infty}{}_n \, c_n \, \vartheta(v - v_n).$$

§ 1.

Eine lineare Funktionalgleichung.

Die Bestimmung analytischer Funktionen kann vollzogen werden auf verschiedene Weise, etwa durch

(1) größenmäßige Darstellung,

(2) Funktionalgleichung,

(3) unabhängige Eigenschaften.

Diese drei Verfahren wünschen wir hinsichtlich ihrer Leistungsfähigkeit zu vergleichen.

Durch Rechenausdrücke (1) wird jedenfalls die Existenz der zu definierenden Funktion einsichtig. Für jeden endlichen Bereich von Argumenten ist damit aber die Funktion auch schon überbestimmt, was ihre Transformierbarkeit hemmt. Diese Hemmung wird fühlbar, weil die Ausgangsdarstellung nicht alle inneren Eigenschaften der Funktion gleichermaßen ins Licht setzen kann; einige erscheinen so verdunkelt, daß man sie aus einem Rechenausdruck nicht direkt entnehmen kann.

Durch Funktionalgleichungen (2) entfällt die Last der Überbestimmung, auch werden wir frei zur zweckmäßigen Auswahl darstellender Elemente in (1) als Folge der Struktur von (2); z. B. pflegt man von Differentialgleichungen aus zu darstellenden Potenzreihen überzugehen, von Differenzengleichungen aber zu Fakultätenreihen. Dagegen ist der Existenzbeweis für Lösungen von (2) nicht trivial und ergibt sich manchmal erst, nachdem die Eigenschaften der Lösung in jedem Teil ihres Existenzgebietes aus (2) entnommen wurden.

Das Verfahren (3), durch Postulierung eines Systems unabhängiger Eigenschaften die gewünschte Funktion zu kennzeichnen, entsteht aus (2) durch einen Aufschließungsprozeß. An Stelle einer identischen Aussage in mehreren Parametern treten verschiedene unabhängige Angaben, die je möglichst inhaltsarm sind. Von selbst lösen sich dabei Größeneigenschaften aus den formelmäßigen Zusammenhängen heraus.

Das Auffinden neuer Eigenschaften wird oft auf dem zuerst skizzierten Wege gelingen; die übersichtliche Zusammenfassung des Tatsachenmaterials läßt später die Benutzung des dritten Verfahrens zweckmäßig erscheinen. Wir verwenden es hier zur Erklärung von Kroneckers s-Funktion.

Fünf komplexe Größen $\omega_1, \omega_2, x, y, u$ seien den vier Bedingungen unterworfen

$$0 < |\omega_1|; \qquad 0 < \left| J\left(\frac{\omega_2}{\omega_1}\right) \right|, \tag{4}$$

$$\omega_1 y - \omega_2 x \not\equiv 0 \pmod{\omega_1, \omega_2}, \tag{5}$$

$$u \not\equiv 0 \pmod{\omega_1, \omega_2}. \tag{6}$$

Eine in x, y, u meromorphe Funktion $s\binom{x\ y}{u} = s(u)$ kann bestimmt werden durch folgende Angaben.

Für jedes zulässige u sei

$$s(u) \tag{7}$$

regulär; es bestehe die Grenzbedingung

$$\lim_{u \to 0} u\, s(u) = 1; \tag{8}$$

im Großen gelte

$$s(u) = e^{2\pi i x} s(u + \omega_1) = e^{2\pi i y} s(u + \omega_2). \tag{9}$$

Nämlich es gibt nicht mehr als eine Lösung für (7), (8), (9), wie durch Liouvilles Prinzip oder Überlegungen betreffend das Wachstum ganzer Transzendenten zu erschließen ist. Die Eindeutigkeit der Lösung verschafft uns eine Funktionalgleichung für $s(u)$, falls es gelingt, einen weiteren Ausdruck zu konstruieren, der (7), (8), (9) erfüllt. Es ist nach (9) für $\nu = 1, 2, \ldots$

$$s\begin{pmatrix} \nu x, \nu y \\ \frac{u}{\nu} \end{pmatrix} = e^{2\pi i \nu x} s\begin{pmatrix} \nu x, \nu y \\ \frac{u}{\nu} + \omega_1 \end{pmatrix}.$$

Bildet man eine ν^2-gliedrige Summe

$$\frac{1}{\nu} \sum_{0}^{\nu-1} \sum_{\beta}^{\alpha} e^{2\pi i(x\alpha + y\beta)} s\begin{pmatrix} \nu x, \nu y \\ \frac{u + \omega_1 \alpha + \omega_2 \beta}{\nu} \end{pmatrix} = \sigma(u), \tag{10}$$

so ist die affine Streckung des Grundgebietes in der u-Ebene durch die Doppelsumme (10) kompensiert, also auch für $\sigma(u)$ (7), (8) erfüllt. Wegen

$$\sigma(u + \omega_1) = \frac{1}{\nu} \sum_{0}^{\nu-1} \sum_{\beta}^{\alpha} e^{2\pi i(x\alpha + y\beta)} s\begin{pmatrix} \nu x, \nu y \\ \frac{u + \omega_1 + \omega_1 \alpha + \omega_2 \beta}{\nu} \end{pmatrix}$$

$$= \frac{1}{\nu} \sum_{0}^{\nu-1} e^{2\pi i y \beta} \sum_{1}^{\nu} e^{2\pi i x(\alpha - 1)} s\begin{pmatrix} \nu x, \nu y \\ \frac{u + \omega_1 \alpha + \omega_2 \beta}{\nu} \end{pmatrix} = e^{-2\pi i x} \sigma(u)$$

gilt aber auch (9), so daß $\sigma = s$ folgt, was eine in s lineare Identität liefert

$$s\begin{pmatrix} x, y \\ u \end{pmatrix} = \frac{1}{\nu} \sum_{0}^{\nu-1} \sum_{\beta}^{\alpha} e^{2\pi i(x\alpha + y\beta)} s\begin{pmatrix} \nu x, \nu y \\ \frac{u + \omega_1 \alpha + \omega_2 \beta}{\nu} \end{pmatrix}. \tag{11}$$

Diese homogene Funktionalgleichung (11) wird auch befriedigt durch jedes konstante Vielfache von s, und kann daher (8) nicht enthalten. Wenn also die x, y, ν enthaltende Aussage (11) dem System (7), (8), (9) nicht gleichwertig ist, so reicht sie für $\frac{\partial c}{\partial u} = 0$ doch aus zur Kennzeichnung von $c\,s(u)$. Sollten in (11) insbesondere vertauschbare Grenzprozesse vorliegen, so daß der Grenzwert

$$\lim_{\nu\to\infty}\frac{1}{\nu}\sum_{0}^{\nu-1}\sum_{\beta}^{\alpha} e^{2\pi i(x\alpha+y\beta)}\, s\left(\frac{\nu x,\ \nu y}{u+\omega_1\alpha+\omega_2\beta}{\nu}\right) = \lim_{\nu\to\infty}\sum_{0}^{\nu-1}\sum_{\beta}^{\alpha}\frac{e^{2\pi i(x\alpha+y\beta)}}{u+\omega_1\alpha+\omega_2\beta}\cdot\lim_{\substack{\nu\to\infty\\ \varepsilon\to 0}}\varepsilon\cdot s\left(\begin{matrix}\nu x,\ \nu y\\ \varepsilon\end{matrix}\right.$$

existiert, so führt obige lineare Funktionalgleichung unmittelbar zu einer trigonometrischen Darstellung; aber die Vertauschbarkeit der obigen Grenzprozesse hat nicht statt, und in (11) ist das Summationsgebiet abzuändern, ehe ein Grenzübergang möglich wird. In Wahrheit gilt nämlich für

(12) $$x, y \not\equiv 0\ (1),$$

(12′) $$s(u) = \lim_{\nu\to\infty}\sum_{-\nu}^{\nu}\sum_{\beta}^{\alpha}\frac{e^{2\pi i(x\alpha+y\beta)}}{u+\omega_1\alpha+\omega_2\beta},$$

und zwar kann (12), (12′) aus (11) oder in unbestimmtem Ansatz aus (7), (8), (9) hergeleitet werden.

§ 2.

Nullstellen der s-Funktion.

Aus § 1 (7), (8), (9) folgt unmittelbar

(1) $$s\left(\begin{matrix}x, y\\ u\end{matrix}\right) + s\left(\begin{matrix}-x, -y\\ -u\end{matrix}\right) = 0.$$

Setzt man also

$$s\left(\begin{matrix}x, y\\ u\end{matrix}\right) = \frac{1}{u} + t(x, y) + O(u),$$

so folgt

(1′) $$t(x, y) + t(-x, -y) = 0.$$

Bildet man

(1″) $$s\left(\begin{matrix}x, y\\ u\end{matrix}\right) s\left(\begin{matrix}-x, -y\\ u\end{matrix}\right) = \varphi\left(\begin{matrix}x, y\\ u\end{matrix}\right),$$

so wird nach § 1 (8)

$$\varphi\left(\begin{matrix}x, y\\ u\end{matrix}\right) = \frac{1}{u^2} + (*) + b(x, y) + O(u),$$

wegen § 1 (9) aber

$$\varphi\left(\begin{matrix}x, y\\ u\end{matrix}\right) = \varphi\left(\begin{matrix}x, y\\ u+\omega_1\end{matrix}\right) = \varphi\left(\begin{matrix}x, y\\ u+\omega_2\end{matrix}\right).$$

Zusammen mit § 1 (7) liefert dies

$$\varphi\begin{pmatrix} x, y \\ u \end{pmatrix} = \wp(u, \omega_1, \omega_2) - c(x, y), \tag{2}$$

führt also auf Weierstraßens elementare doppelperiodische Funktion $\wp(u, \omega_1, \omega_2) = \wp(u)$. Um in (2) die von u freie Größe $c(x, y)$ näher zu bestimmen, greifen wir nochmals auf die Definition von $s\begin{pmatrix} x, y \\ u \end{pmatrix}$ zurück. Werden x, y um ganze Zahlen verändert, so bleiben § 1 (7), (8) unverändert, während auf § 1 (9) das Additionstheorem der Exponentialfunktion anzuwenden ist; man entnimmt

$$s\begin{pmatrix} x, y \\ u \end{pmatrix} = s\begin{pmatrix} x+1, y \\ u \end{pmatrix} = s\begin{pmatrix} x, y+1 \\ u \end{pmatrix}. \tag{3}$$

Die Periodizität (3) reicht keineswegs aus, um die Funktion in Abhängigkeit ihrer oberen Argumente zu kennen. Wir haben also weitere Eigenschaften in x, y herzuleiten. Für komplexes v setze man $s\begin{pmatrix} x+\omega_1 v, y+\omega_2 v \\ u \end{pmatrix} = \psi(u)$; dies genügt den Forderungen § 1 (7), (8), während § 1 (9) zu ersetzen ist durch

$$\psi(u) = e^{2\pi i(x+\omega_1 v)}\psi(u+\omega_1) = e^{2\pi i(y+\omega_2 v)}\psi(u+\omega_2).$$

Andererseits betrachte man

$$e^{-2\pi i u v}\, s\begin{pmatrix} x, y \\ u \end{pmatrix} = \chi(u),$$

wofür wegen $e^0 = 1$ ebenfalls § 1 (7), (8) erfüllt ist; hier entsteht

$$\chi(u+\omega_1) = e^{-2\pi i \omega_1 v} e^{-2\pi i u v}\, s\begin{pmatrix} x, y \\ u+\omega_1 \end{pmatrix}$$

und nach § 1 (9)

$$= e^{-2\pi i(x+\omega_1 v)}\chi(u).$$

Derselbe Prozeß auf y und ω_2 angewandt, zeigt

$$\chi(u) = e^{2\pi i(x-\omega_1 v)}\chi(u+\omega_1) = e^{2\pi i(y+\omega_2 v)}\chi(u+\omega_2),$$

und es ist identisch in x, y, u, v stets $\psi = \chi$ oder

$$s\begin{pmatrix} x+\omega_1 v, y+\omega_2 v \\ u \end{pmatrix} = e^{-2\pi i u v}\, s\begin{pmatrix} x, y \\ u \end{pmatrix}. \tag{4}$$

Die vorliegende Beziehung verwerten wir für $\varphi(x, y)$ in (2); tauscht man in (4) x, y, v gegen $-x, -y, -v$, so kann durch Produktbildung der Exponentialfaktor entfernt werden. Man erhält

$$\varphi\begin{pmatrix} x, y \\ u \end{pmatrix} = \varphi\begin{pmatrix} x+\omega_1 v, y+\omega_2 v \\ u \end{pmatrix}. \tag{5}$$

Für die in § 1 (5) auftretende Größe werde als abkürzende Bezeichnung

$$\omega_1 y - \omega_2 x = z \tag{6}$$

eingeführt; der durch (5) gegebene Ausdruck zeigt sich dann als analytisch abhängig von u und z; die gleiche Reduktion tritt in (4) ein für $v = -\frac{x}{\omega_1}$, wo

$$e^{\frac{2\pi i \chi u}{\omega_1}} s\begin{pmatrix} x, y \\ u \end{pmatrix} = s\begin{pmatrix} 0, y - \frac{\omega_2}{\omega_1} x \\ u \end{pmatrix} = s\begin{pmatrix} 0, \frac{z}{\omega_1} \\ u \end{pmatrix}$$

beschrieben werden kann in nur zwei Veränderlichen. Setzt man

$$e^{\frac{2\pi i \chi u}{\omega_1}} s\begin{pmatrix} x, y \\ u \end{pmatrix} = f(u, z), \tag{7}$$

so zeigt § 1 (7), (8), (9) für $f(u)$ folgendes kennzeichnende Verhalten nahe der Stelle $u = 0$

$$f(u) = f(u + \omega_1) = e^{\frac{2\pi i z}{\omega_1}} f(u + \omega_2) = \frac{1}{u} + \dots .$$

Cauchysche Integration längs eines Parallelogrammes ergibt wegen der Abwesenheit störender Pole für $\alpha = -\frac{\omega_2}{2\omega_1}$

$$\int_\alpha^{\alpha+\omega_1} du f(u, z) = \frac{2\pi i}{-e^{\frac{-2\pi i z}{\omega_1}} + 1} = \frac{\omega_1}{z} + \dots , \tag{7'}$$

woraus wegen (7) der Grenzwert fließt

$$\lim_{z \to 0} z s\begin{pmatrix} x, y \\ u \end{pmatrix} = 1. \tag{8}$$

Durch (3), (4), (8) ist s in Abhängigkeit von x und y bestimmt. Setzt man in Ergänzung von (2), (5), (6)

$$c(x, y) = a(z),$$

so gibt (1″), (8)

$$\lim_{z \leftarrow 0} z^2 \cdot a(z) = 1, \tag{9}$$

während die analytische Funktion $a(z)$ vermöge (3) sich elliptisch erweist

$$a(z) = a(z + \omega_1) = a(z + \omega_2). \tag{9'}$$

Die Abwesenheit weiterer Pole in einem Parallelogramm nahe der Stelle $z = 0$ gibt

$$a(z) = \wp(z),$$

weshalb (2) endgültig die Gestalt annimmt

$$s\begin{pmatrix} x, y \\ u \end{pmatrix} \cdot s\begin{pmatrix} -x, -y \\ u \end{pmatrix} = \wp(u) - \wp(z). \tag{10}$$

Die Differenz von $\wp$-Funktionen rechts wird durch (10) in s-Faktoren aufgespalten; dies entspricht der Tatsache, daß in der Teilbruchreihe für $\wp(u)$ die Teilnenner vom Grade zwei, jedoch für $s(u)$ nur vom ersten Grad sind.

Verständlich wird, daß an Stelle des Additionstheoremes vierten Grades, welches für $\wp(u)$ statthat, eine Aussage niedrigeren Grades in $s\binom{x,y}{u}$ tritt. Als nächste Folgerung aus (10) entnehmen wir aber die Nullstellen von $s\binom{x,y}{u}$. Da die vier Ecken $0, \omega_1, \omega_2, (\omega_1+\omega_2)$ eines Grundparallelogrammes Pole erster Ordnung in der u-Ebene sind, kann innerhalb jenes Parallelogrammes nur eine Nullstelle $u=u_0$ liegen. Aus (10) folgt

$$u_0^2 = z^2;$$

weil jedoch nur einer der Faktoren $s\binom{x,y}{z}$, $s\binom{-x,-y}{z}$ schwindet, genügt es, den nichtverschwindenden in einem Sonderfalle festzustellen, etwa für $z \to 0$. Aus § 1 (8), § 2 (8) entnimmt man

$$\lim_{z\to 0} z\, s\binom{x,y}{z} = 2,$$

woraus

$$s\binom{x,y}{z} \neq 0$$

und identisch in x, y

$$s\binom{x,y}{-z} = 0 \tag{11}$$

folgt, und die Nullstelle

$$u_0 = -z.$$

§ 3.

Existenzbeweis für $s\binom{x,y}{u}$.

Die durch § 2 (7) eingeführte Funktion war in Abhängigkeit von u gekennzeichnet durch

$$f(u) = f(u+\omega_1) = e^{\frac{2\pi i z}{\omega_1}} f(u+\omega_2) = \frac{1}{u} + \dots . \tag{1}$$

Ihre Abhängigkeit vom zweiten Argument z war durch § 2 (3), (7), (7′) festgelegt und ergibt

$$f(u,z) = f(u, z+\omega_1) = e^{\frac{2\pi i u}{\omega_1}} f(u, z+\omega_2) = \frac{1}{z} + \dots . \tag{2}$$

Vergleicht man (1) und (2), so folgt Symmetrie hinsichtlich beider Argumente

$$f(u,z) = f(z,u)$$

und nach § 2 (7), (7′), wenn $-\frac{\omega_2}{2\omega_1} = \alpha$ gesetzt wird,

$$\frac{1}{2\pi i}\int_{\alpha}^{\alpha+\omega_1} dz\, e^{\frac{2\pi i z u}{\omega_1}}\, s\binom{x,y}{u} = \frac{1}{-e^{-2\pi i u}+1}. \tag{3}$$

Für $\begin{cases} 0 < x, y < 1 \\ h, k \equiv 0 \ (1) \end{cases}$ berechnet sich aus (3) sehr einfach:

$$\int_0^1\int dx\,dy\, e^{-2\pi i(xh+yk)}\, s\begin{pmatrix} x, y \\ u \end{pmatrix} = \int_0^1 dx\, e^{-2\pi i x\left(\frac{u}{\omega_1}+h+\frac{\omega_2}{\omega_1}k\right)} \cdot \frac{1}{\omega_1}\int_0^{\alpha+\omega_1} dz\, e^{-\frac{2\pi i}{\omega_1}(zk-xu)}\, s\begin{pmatrix} x, y \\ u \end{pmatrix}$$

$$= \frac{1}{\omega_1}\int_0^1 dx\, e^{-\frac{2\pi i}{\omega_1}x(u+\omega_1 h+\omega_2 k)} \cdot \int_\alpha^{\alpha+\omega_1} dz\, e^{\frac{2\pi i}{\omega_1}zk} \cdot e^{\frac{2\pi i x u}{\omega_1}}\, s\begin{pmatrix} x, y \\ u \end{pmatrix}$$

$$= \frac{c\, e^{-\frac{2\pi i}{\omega_1}(u+h\omega_1+k\omega_2)} - 1}{-2\pi i(u+h\omega_1+k\omega_2)} \cdot \int_\alpha^{\alpha+\omega_1} dz\, e^{-\frac{2\pi i}{\omega_1}zk} f(u,z) \quad \text{(nach (1))}$$

$$= \frac{1-e^{-\frac{2\pi i}{\omega_1}(u+k\omega_2)}}{u+h\omega_1+k\omega_2} \cdot \frac{1}{2\pi i}\int_\alpha^{\alpha+\omega_1} dz\, f(u+k\omega_2, z).$$

Wegen § 2 (7), § 3 (3) entsteht daher endgültig

$$(4) \qquad \int_0^1\int dx\,dy\, e^{-2\pi i(xh+yk)}\, s\begin{pmatrix} x, y \\ u \end{pmatrix} = \frac{1}{u+\omega_1 h+\omega_2 k}.$$

Die Integralformel (4) leistet die strenge Begründung der in § 1 (12′) angegebenen trigonometrischen Reihe

$$(5) \qquad s(u) = \lim_{\nu\to\infty} \sum_{-\nu}^{\nu}{}^{h}_{k} \frac{e^{2\pi i(xh+yk)}}{u+\omega_1 h+\omega_2 k} \quad \text{für} \quad 0 < x, y < 1.$$

Daß nicht nur der Grenzwert (5) existiert, sondern auch zusammenfällt mit anders erstreckten Summen, wie

$$\lim_{\nu\to\infty} \sum_{k}^{h}{}^{h^2+k^2<\nu} \frac{e^{2\pi i(xh+yk)}}{u+\omega_1 h+\omega_2 k},$$

folgt aus Cauchys Integralsatz und § 1 (7), (8), (9) sofort. Wegen § 2 (3) sind für x, y sogar alle nicht ganzen reellen Werte zulässig, wie in § 1 (12), (12′) angekündigt.

In den Paragraphen 1 und 2 wurden Eigenschaften der Funktion bewiesen, „falls $s\begin{pmatrix} x, y \\ u \end{pmatrix}$ existiert“. Durch die Darstellung (5) ist der gewünschte Existenzbeweis geleistet, insbesondere also die Widerspruchsfreiheit der in § 1 (7), (8), (9) enthaltenen vier Postulate dargetan.

Die zahlenmäßige Berechnung von $s\begin{pmatrix} x, y \\ u \end{pmatrix}$ wird freilich durch (5) nicht gefördert; indem wir uns nach besser konvergenten Entwickelungen umsehen, wollen wir die Einheit der Theorie nicht aus dem Auge verlieren. Etwa benötigte Hilfsfunktionen sollten nicht unabhängig definiert werden, sondern aus s-Funktionen heraus konstruierbar sein.

§ 4.

Einführung von $\vartheta\left(\frac{z}{\omega_1}\right)$.

Die praktische Ausrechnung von $s\binom{x,\,y}{u}$ macht eine Umgestaltung des Ausdruckes § 3 (5) im Sinne besserer Konvergenz erwünscht; wir werden eine Hilfsfunktion beschaffen, die nicht von x, y, u abhängt, sondern nur von einem einzigen Argument, etwa $\omega_1 y - \omega_2 x = z$. Die gewünschte Elimination ist schon vorbereitet durch den in § 2 (7) eingeführten Ausdruck

$$e^{\frac{2\pi i x u}{\omega_1}} s\binom{x,\,y}{u} = f(u,z). \tag{1}$$

Die in u meromorphe Funktion (1) hat für $u = 0$ einen Pol, nach § 2 (11) für $u = -z$ eine Nullstelle, beide von erster Ordnung; sonst aber in dem durch ω_1, ω_2 aufgespannten Parallelogramm einen endlichen nicht verschwindenden Wert. Allgemein hat nun $f(u+z, \zeta)$ einen kompensierenden Pol in $u = -z$, und insbesondere die Annahme $\zeta = -z$ läßt die Funktion

$$f(u,z)\cdot f(u+z,-z) = h(z) \tag{2}$$

wegen (1) und § 1 (9) innerhalb der ganzen u-Ebene von Nullstellen und Polen frei sein, woraus leicht $h(z)$ als unabhängig von u erkannt wird.

In der z-Ebene spanne man ebenfalls das Parallelogramm ω_1, ω_2 auf, findet nach § 2 (8) nur einen Pol zweiter Ordnung der meromorphen Funktion $h(z)$ darin, und keine Nullstellen, so daß nach § 3 (2) die ganze Transzendente

$$[-h(z)]^{-\frac{1}{2}} = z + \ldots \tag{3}$$

eindeutig gekennzeichnet werden kann. Vermöge

$$\begin{cases} f(u,z) = f(u, z+\omega_1), \\ f(u+z,-z) = f(u+z+\omega_1, -z-\omega_1), \end{cases}$$

$$\begin{cases} f(u,z) = e^{\frac{2\pi i}{\omega_1} u} f(u, z+\omega_2), \\ f(u+z,-z) = e^{-\frac{2\pi i}{\omega_1}(u+2z+\omega_2)} f(u+z+\omega_2, -z-\omega_2) \end{cases}$$

wird

$$h(z) = h(z+\omega_1) = e^{-\frac{2\pi i}{\omega_1}(2z+\omega_2)} h(z+\omega_2). \tag{4}$$

Um (3), (4) zusammenzufassen, führe man die übliche Bezeichnung für Jacobis Transzendente ein und setze

$$[-h(z)]^{-\frac{1}{2}} = \frac{\omega_1 \vartheta\left(\frac{z}{\omega_1}\right)}{\vartheta'(0)}. \tag{5}$$

Aus (2) fließt $h(z) = h(-z)$; aus (3), (5) daher

$$\vartheta\left(\frac{z}{\omega_1}\right) = -\vartheta\left(-\frac{z}{\omega_1}\right), \tag{6}$$

aus (4) alsdann

$$\vartheta\left(\frac{z}{\omega_1}\right) = -\vartheta\left(\frac{z}{\omega_1}+1\right) = -e^{\frac{\pi i}{\omega_1}(2z+\omega_2)}\,\vartheta\left(\frac{z+\omega_1}{\omega_1}\right) \tag{6'}$$

und wegen (5) ist

$$\lim_{z\to\infty}\frac{\omega_1\,\vartheta\left(\frac{z}{\omega_1}\right)}{z\,\vartheta'(0)} = 1. \tag{6''}$$

Mittels (6), (6'), (6'') ist nicht die ungerade Thetafunktion selbst definiert, sondern nur der Quotient $\frac{\vartheta\left(\frac{z}{\omega_1}\right)}{\vartheta'(0)}$; in der That besteht noch immer betreffend die Perioden die in § 1 (4) gesetzte Einschränkung

$$0 < J\left(\frac{\omega_2}{\omega_1}\right) \tag{7}$$

und nicht die stärkere Bindung, welche beim üblichen Durchgang durch Modulfunktionen zu fordern ist,

$$0 < J\left(\frac{\omega_2}{\omega_1}\right). \tag{7'}$$

Während Zähler und Nenner des Quotienten (5) einzeln der Ungleichung (7') unterworfen sind, ist $[h(z)]^{-\frac{1}{2}}$ selbst schon unter der Annahme (7) vorhanden; der Grund dafür ist, daß ein Zeichenumschlag von ω_2 auch $z = \omega_1 y - \omega_2 x$ verändert.

Nunmehr soll (7') statthaben; wegen $\left|e^{\pi i\frac{\omega_2}{\omega_1}}\right| < 1$ konvergiert $\sum_{-\infty}^{+\infty} e^{\pi i\frac{\omega_2}{\omega_1}(\nu+\frac{1}{2})^2}$, und die aus (5), (6'), (6'') sich ergebende Darstellung

$$\frac{\omega_1\,\vartheta\left(\frac{z}{\omega_1}\right)}{\vartheta'(0)} = \frac{\sum_{-\infty}^{\infty}\nu\,(-1)^\nu\, e^{\pi i\left[(\nu+\frac{1}{2})^2\frac{\omega_2}{\omega_1}+\frac{2z}{\omega_1}(\nu+\frac{1}{2})\right]}}{\sum_{-\infty}^{\infty}\nu\,(-1)^\nu\,(2\nu+1)\,e^{\pi i(\nu+\frac{1}{2})^2\frac{\omega_2}{\omega_1}}} \tag{8}$$

erlaubt bei vorgegebenen $\omega_1, \omega_2, x, y, u$ in wenigen Schritten

$$f(u,z) = \frac{\vartheta'(0)\,\vartheta\left(\frac{u+z}{\omega_1}\right)}{\omega_1\,\vartheta\left(\frac{u}{\omega_1}\right)\vartheta\left(\frac{z}{\omega_1}\right)}$$

und daraus

$$s\begin{pmatrix}x,y\\u\end{pmatrix} = e^{-\frac{2\pi i x u}{\omega_1}}\cdot\frac{\vartheta'(0)\,\vartheta\left(\frac{u+z}{\omega_1}\right)}{\omega_1\,\vartheta\left(\frac{u}{\omega_1}\right)\vartheta\left(\frac{z}{\omega_1}\right)} \tag{9}$$

beliebig genau zu berechnen.

Durch (8), (9) ist die praktische Ausrechnung von $s\binom{x,y}{u}$ geleistet; umgekehrt zeigen (1), (2), (5), daß man durch Spezialisierung der s-Argumente auf rationalem Wege nicht nur zur $\wp$- sondern auch zur ϑ-Funktion geführt wird. Es ist

$$\vartheta\left(\frac{z}{\omega_1}\right) = \frac{\vartheta'(0)\, e^{\pi i \frac{xz}{\omega_1}}}{\omega_1 \sqrt{s\binom{x,y}{u}\, s\binom{x,y}{-u,-z}}}$$

und insbesondere

$$\vartheta\left(\frac{z}{\omega_1}\right) = \frac{\vartheta'(0)}{\omega_1} \frac{e^{\frac{\pi i x z}{\omega_1}}}{s\binom{x\ y}{-\frac{z}{2}}}.$$

§ 5.

Der Additionssatz.

Die für $s\binom{x,y}{u} = s(u)$ in § 1 (11) aufgestellte lineare Identität kann mit Hilfe von § 1 (7), (8), (9) verallgemeinert werden. Für $\nu = 1, 2, \ldots$ bezeichne das Symbol $\sum_{\beta}^{\alpha \bmod \nu}$ eine Summation, worin jeder Zeiger unabhängig ein volles Restsystem mod ν durchläuft. Es sei $0 < |\omega_1|$; $0 < J\left(\frac{\omega_2}{\omega_1}\right)$; $\omega_1 y - \omega_2 x \not\equiv 0 \bmod \omega_1, \omega_2$; $u \not\equiv 0 \bmod \omega_1, \omega_2$; $\nu = 1, 2, \ldots$. Dann gilt

$$(1) \qquad s(u) = \frac{1}{\nu} \sum_{\beta}^{\alpha \bmod \nu} e^{2\pi i(x\alpha + y\beta)}\, s\left(\frac{\nu x,\ \nu y}{u + \omega_1 \alpha + \omega_2 \beta}\right).$$

Aus (1) entnimmt man unmittelbar § 1 (7), (9); dagegen enthält die homogene Aussage (1) nicht die Grenzbedingung § 1 (8), sondern reicht nur aus zur Kennzeichnung von $c \cdot s(u)$ mit $\frac{\partial c}{\partial u} = 0$.

Eine Lösung von (1), die mit $u \to 0$ endlich bleibt, läßt daher auf $c = 0$ schließen: sie verschwindet identisch in u.

Der vorliegende Gedanke überträgt sich auf Funktionen zweier Veränderlicher $\sigma(u_1, u_2)$. Es bezeichne das Symbol $\sum_{\beta_1, \beta_2}^{\alpha_1, \alpha_2 \bmod \nu}$ eine vierfache Summe, darin jeder Zeiger ein volles Restsystem mod ν durchläuft. Es seien $u_1, u_2 \not\equiv 0 \bmod \omega_1, \omega_2$ und x, y, ξ, η beliebig; besteht dann

$$(2) \qquad \overline{\lim_{u_1, u_2 \to 0}} |\sigma(u_1, u_2)| < \infty$$

zusammen mit

$$(3) \qquad \sigma\binom{x, y, \xi, \eta}{u_1, u_2} = \frac{1}{\nu^2} \sum_{\beta_1 \beta_2}^{\alpha_1, \alpha_2 \bmod \nu} e^{2\pi i(x\alpha_1 + y\beta_1 + \xi\alpha_2 + \eta\beta_2)}\, \sigma\left(\frac{\nu x,\ \nu y,\ \nu \xi,\ \nu \eta}{\frac{u_1 + \omega_1 \alpha + \omega_2 \beta}{\nu}, \frac{u_2 + \omega_1 \alpha + \omega_2 \beta}{\nu}}\right),$$

so folgt wieder aus (3), daß die Stelle $u_1 = 0,\ u_2 = 0$ zu den Singularitäten von $\sigma(u_1, u_2)$ gehört; und wegen seiner Beschränktheit muß σ identisch verschwinden.

Nach dieser Vorbereitung gehen wir daran, die drei Größen ν, ω_1, ω_2 aus der linearen Funktionalgleichung (1) zu entfernen. Dazu sind drei weitere Argumente v, ξ, η mit $v \not\equiv 0;\ \omega_1\eta - \omega_2\xi \not\equiv 0 \bmod \omega_1, \omega_2$ einzuführen. Die in Argumenten und Zeigern bilineare Form

$$(4)\quad x\alpha_1 + y\beta_1 + \xi\alpha_2 + \eta\beta_2 = x(\alpha_1 - \alpha_2) + y(\beta_1 - \beta_2) + (x+\xi)\alpha_2 + (y+\eta)\beta_2$$

$$(4')\quad = (x+\xi)\alpha_1 + (y+\eta)\beta_1 + \xi(\alpha_2 - \alpha_1) + \eta(\beta_2 - \beta_1)$$

führt zur Bildung von s-Produkten, die (4) enthalten, wie

$$(5)\quad s\begin{pmatrix} x, y \\ u \end{pmatrix} s\begin{pmatrix} \xi, \eta \\ v \end{pmatrix} = \frac{1}{\nu^2} \sum_{\beta_1\beta_2}^{\mathrm{mod}\,\nu} {}^{\alpha_1\alpha_2} e^{2\pi i(x\alpha_1 + y\beta_1 + \xi\alpha_2 + \eta\beta_2)} s\left(\frac{\nu x,\ \nu y}{u + \omega_1\alpha + \omega_2\beta \over \nu}\right) s\left(\frac{\nu\xi,\ \nu\eta}{v + \omega_1\alpha + \omega_2 \over \nu}\right)$$

Wendet man auf (5) die Umformung (4) an, ordnet um und ersetzt u durch $u - v$ in den s-Argumenten sowie ξ, η durch $x + \xi,\ y + \eta$, so entsteht entsprechend

$$(5'')\quad s\begin{pmatrix} x, y \\ u - v \end{pmatrix} s\begin{pmatrix} x+\xi, y+\eta \\ v \end{pmatrix}$$

$$= \frac{1}{\nu^2} \sum_{\beta_1\beta_2}^{\mathrm{mod}\,\nu} {}^{\alpha_1\alpha_2} e^{2\pi i(x\alpha_1 + y\beta_1 + \xi\alpha_2 + \eta\beta_2)} s\left(\frac{\nu x,\ \nu y}{u + \omega_1\alpha_1 + \omega_2\beta_1 \over \nu}\right) s\left(\frac{\nu x + \nu\xi,\ \nu y + }{v + \omega_1\alpha_2 + \omega_2 \over \nu}\right)$$

Ähnlich ist wegen (4′) noch

$$(5'')\quad s\begin{pmatrix} x+\xi, y+\eta \\ u \end{pmatrix} s\begin{pmatrix} \xi\ \eta \\ v - u \end{pmatrix}$$

zu behandeln; kombiniert man (5), (5′), (5″) mit konstanten $\varkappa, \lambda$ zu

$$s\begin{pmatrix} x, y \\ u \end{pmatrix} s\begin{pmatrix} \xi, \eta \\ v \end{pmatrix} - \varkappa s\begin{pmatrix} x, y \\ u - v \end{pmatrix} s\begin{pmatrix} x+\xi, y+\eta \\ v \end{pmatrix} - \lambda s\begin{pmatrix} \xi, \eta \\ v - u \end{pmatrix} s\begin{pmatrix} x+\xi, y+\eta \\ u \end{pmatrix} = \sigma_{\varkappa\lambda}(u, v)$$

so genügt $\sigma_{\varkappa\lambda}(u_1, u_2)$ der Identität (3). Ein Vergleich mit

$$\frac{1}{u - v} - \frac{1}{(u-v)v} - \frac{1}{(v-u)u} = 0$$

zeigt, daß insbesondere $\sigma_{1,1}(u_1, u_2)$ der Ungleichung (2) genügt, weshalb $\sigma_{1,1}(u, v) = 0$ wird. Somit entsteht für

$$u - v \not\equiv 0, \qquad z + \omega_1\eta - \omega_2\xi \not\equiv 0 \pmod{\omega_1, \omega_2}$$

die gewünschte quadratische Funktionalgleichung der s-Funktion

$$(6)\quad s\begin{pmatrix} x, y \\ u \end{pmatrix} s\begin{pmatrix} \xi, \eta \\ v \end{pmatrix} = s\begin{pmatrix} x, y \\ u - v \end{pmatrix} s\begin{pmatrix} x+\xi, y+\eta \\ v \end{pmatrix} + s\begin{pmatrix} \xi, \eta \\ v - u \end{pmatrix} s\begin{pmatrix} x+\xi, y+\eta \\ u \end{pmatrix}.$$

Seien a, b unabhängig von x, y, ξ, η, u, v; mit $s\binom{x,y}{u}$ genügt dann auch die allgemeinere Funktion

$$e^{a+bxu}\, s\binom{x,y}{u} = \sigma\binom{x,y}{u} \tag{7}$$

der Identität (6). Daß aber durch (7) das volle Lösungssystem für (6) geliefert wird, erkennt man folgendermaßen: Trägt man § 4 (9) in (6) ein, so findet man Weierstraßens „dreigliedrige Beziehung“ für die ϑ-Funktion. Nach Hermite und Hurwitz kennzeichnet diese „équation à trois termes“ die Funktionenschar $e^{c+dz^2}\,\vartheta\left(\frac{z}{\omega_1}\right)$ mit konstanten c, d; wegen § 4 (9) schließt man direkt auf (7).

Damit ist der Gehalt von (6) noch nicht erschöpft. Ersetzt man u, v durch $-u, -v$, verwendet § 2 (10), benutzt die Nullstellen von $s\binom{x,y}{u}$ und entfernt s zugunsten von $\wp$, so gelangt man zum Additionstheorem der $\wp$-Funktion. Mit $\omega_1+\omega_2+\omega_3 = u_1+u_2+u_3 = 0$, $\wp\left(\frac{\omega_j}{2}\right) = e_j$, $\wp(u_j) = \wp_j$ lautet dies in rationaler Gestalt

$$\begin{aligned} &4\,(\wp_1+\wp_2+\wp_3)\,[\wp_1\wp_2\wp_3 - e_1e_2e_3] \\ &\qquad = \{\wp_1\wp_2+\wp_2\wp_3+\wp_3\wp_1 - e_1e_2 - e_2e_3 - e_3e_1\}^2; \end{aligned}$$

dasselbe steht im Mittelpunkt der Weierstraßschen Theorie.

§ 6.

Orthogonalität.

Als „Additionssatz“ war eine quadratische Funktionalgleichung eingeführt worden, worin mehrere Argumentpaare additiv oder subtraktiv verbunden waren. Wir deuteten an, wie daraus auf finitem Wege ein „Additionstheorem“ gewonnen wird, welches zur Kennzeichnung von $\wp(z)$ ausreicht. Durch einen Grenzprozeß finden wir aus der nämlichen Quelle Differentialbeziehungen, die für die s-Funktionen Erscheinungen ins Licht setzen, welche der klassischen $\wp$-Theorie fremd sind. Analog dem elementaren Grenzfall der Exponentialfunktion können auch hier orthogonale Funktionensysteme konstruiert werden, eine Tatsache, die zur Entwicklung analytischer Funktionen nach elliptischen führt.

Man schreibe

$$s\binom{x,y}{u} = s\binom{x}{u}; \qquad s\binom{\xi,\eta}{u} = s\binom{\xi}{u}$$

und erkläre durch

$$\frac{1}{2\pi i}\oint_{0+}\frac{du}{u}\, s\binom{x}{u} = t(x) \tag{1}$$

eine von u freie Größe, dann gibt § 5 (2) für $v \to u$

$$s\binom{x}{u}\, s\binom{\xi}{u} = [t(x)+t(\xi)]\, s\binom{x+\xi}{u} - s'\binom{x+\xi}{u}. \tag{2}$$

Aus (1) und § 4 (9) folgt für $z \not\equiv 0 \pmod{\omega_1, \omega_2}$

$$t(x) = -2\pi i \frac{x}{\omega_1} + \frac{1}{\omega_1} \frac{\vartheta'}{\vartheta}\left(\frac{z}{\omega_1}\right). \tag{1'}$$

Dann gilt

$$\oint_{0+} du\, e^{-u[t(x)+t(\xi)]}\, s\begin{pmatrix} x \\ u \end{pmatrix} s\begin{pmatrix} \xi \\ u \end{pmatrix} = 0,$$

und deshalb existiert unabhängig vom Integrationsweg für $\alpha \not\equiv 0 \pmod{\omega_1, \omega_2}$

$$\int_{\alpha}^{\alpha+\omega_1} du\, e^{-u[t(x)+t(\xi)]}\, s\begin{pmatrix} x \\ u \end{pmatrix} s\begin{pmatrix} \xi \\ u \end{pmatrix} = \left\{ s\begin{pmatrix} x+\xi \\ u \end{pmatrix} e^{-u[t(x)+t(\xi)]} \right\}_{\alpha+\omega_1}^{\alpha} \tag{3}$$

$$= s\begin{pmatrix} x+\xi \\ \alpha \end{pmatrix} e^{-\alpha[t(x)+t(\xi)]} \left\{ 1 - e^{-2\pi i(x+\xi)-\omega_1[t(x)+t(\xi)]} \right\}.$$

Die rechte Seite in (3) zeigt, daß die Bedingungen

$$x + \xi \not\equiv 0\,(1), \tag{4}$$

$$x + \xi + \frac{\omega_1}{2\pi i}[t(x) + t(\xi) \equiv 0\,(1) \tag{4'}$$

für das Verschwinden des Integrales ausreichen:

$$\int_{\alpha}^{\alpha+\omega_1} du\, e^{-u[t(x)+t(\xi)]}\, s\begin{pmatrix} x \\ u \end{pmatrix} s\begin{pmatrix} \xi \\ u \end{pmatrix} = 0. \tag{5}$$

Die Aussagen (4), (4'), (5) vermitteln die Orthogonalität des Funktionensystemes

$$e^{-u\,t(x)}\, s\begin{pmatrix} x \\ u \end{pmatrix} = e^{-\frac{2\pi i u}{\omega_1}\left[x + \frac{\omega_1}{2\pi i} t(x)\right]} \frac{\vartheta(0)\,\vartheta\begin{pmatrix} u+z \\ \omega_1 \end{pmatrix}}{\omega_1 \vartheta\begin{pmatrix} u \\ \omega_1 \end{pmatrix} \vartheta\left(\frac{z}{\omega_1}\right)} = e^{-\frac{2\pi i u}{\omega_1}\left[x + \frac{\omega_1}{2\pi i} t(x)\right]} f(u, z) \tag{6}$$

nach § 4 (9). Wegen (1') kann (4') ersetzt werden durch

$$\frac{1}{2\pi i}\left[\frac{\vartheta'}{\vartheta}\left(y - \frac{\omega_2}{\omega_1} x\right) + \frac{\vartheta'}{\vartheta}\left(\eta - \frac{\omega_2}{\omega_1}\xi\right)\right] \equiv 0\,(1),$$

eine Bedingung, welche jedenfalls befriedigt wird durch

$$\frac{1}{2\pi i} \frac{\vartheta'}{\vartheta}\left(\eta - \frac{\omega_2}{\omega_1}\xi\right) \equiv 0\,(1), \tag{7}$$

$$\frac{1}{2\pi i} \frac{\vartheta'}{\vartheta}\left(y - \frac{\omega_2}{\omega_1} x\right) = \frac{1}{2\pi i} \frac{\vartheta'}{\vartheta}\left(\frac{z}{\omega_1}\right) \equiv 0\,(1). \tag{7'}$$

Aus § 4 (6') fließt $\frac{\vartheta'}{\vartheta}\left(\frac{z}{\omega_1}\right) = \frac{\vartheta'}{\vartheta}\left(\frac{z}{\omega_1} + \frac{\omega_2}{\omega_1}\right) + 2\pi i$, so daß wir an Stelle der Kongruenzen (7), (7') die eine transzendente Bestimmungsgleichung treten lassen können

$$\vartheta'\left(\frac{z}{\omega_1}\right) = 0. \tag{8}$$

Es seien die zugehörigen und mod ω_1 inkongruenten Nullstellen in irgendwelcher Anordnung durch z_ν mit $\nu \equiv 0\,(1)$ bezeichnet. Wegen (4), (5), (6) ist dann zwar

$$(5') \qquad \int_\alpha^{\alpha+\omega_1} du\, f(u, z_\nu)\, f(u, -z_\nu) \neq 0,$$

wohl aber für $\varkappa, \lambda \equiv 0;\ z_\varkappa + z_\lambda \not\equiv 0\,(\omega_2)$

$$(5'') \qquad \int_\alpha^{\alpha+\omega_1} du\, f(u, z_\varkappa)\, f(u, z_\lambda) = 0.$$

Bezeichnet man den Thetaquotienten mit $z = z_\varkappa$

$$\frac{\vartheta'(0)\,\vartheta\left(\frac{u+z_\varkappa}{\omega_1}\right)}{\omega_1\,\vartheta\left(\frac{u}{\omega_1}\right)\vartheta\left(\frac{z_\varkappa}{\omega_1}\right)} = f(u, z_\varkappa) = f_\varkappa,$$

so schreibt sich die Orthogonalitätsrelation (5″) in der Gestalt

$$(9) \qquad \int_\alpha^{\alpha+\omega_1} du\, f_\varkappa f_\lambda = 0.$$

Willkürlich war die Wahl der Weglänge $|\omega_1|$ des Integrationsweges; die Transformation auf beliebige primitive Perioden ist einsichtig. Der festgesetzte besondere Weg und Anfangspunkt α soll jedoch einstweilen beibehalten werden.

§ 7.

Entwicklung nach Eigenfunktionen.

Um die Eigenwerte unseres orthogonalen Funktionensystems leichter zu überblicken, machen wir die Einschränkung

$$(1) \qquad \Re\left(\frac{\omega_2}{\omega_1}\right) = 0,$$

finden für $s(u)$ also ein rechteckiges Grundgebiet. Ebenso bilden in der z-Ebene die Nullstellen von $\vartheta\left(\frac{z}{\omega_1}\right)$ ein Rechtecksgitter. Aus Gründen der Periodizität genügt es, unter den abzählbar vielen Punkten

$$\vartheta'\left(\frac{z}{\omega_1}\right) = 0$$

diejenigen im Vertikalstreifen

$$0 \leqq \Re\left(\frac{z}{\omega_1}\right) < 1$$

herauszugreifen; sie lagern trivialerweise alle entweder auf

$$\Re\left(\frac{z}{\omega_1}\right) = 0$$

oder auf

$$\Re\left(\frac{z}{\omega_1}\right) = \frac{1}{2}.$$

Ordnen wir sie vom Anfangswert $z_0 = \frac{\omega_1}{2}$ nach wachsenden $J\left(\frac{z}{\omega_1}\right)$, so zeigt sich, daß sogar alle z_ν für $\nu \neq 0$ auf der einen Vertikalen

$$\Re\left(\frac{z_\nu}{\omega_1}\right) = 0$$

zu finden sind, übrigens symmetrisch bezüglich der reellen Achse mit

$$z_\nu + z_{-\nu} = 0.$$

Die asymptotische Verteilung der z_ν für große $|\nu|$ ist aus § 4 (6′) durch Differentiation ersichtlich; wegen

$$\frac{\vartheta'}{\vartheta}\left(\frac{z}{\omega_1}\right) = \frac{\vartheta'}{\vartheta}\left(\frac{z}{\omega_1} + \nu\frac{\omega_2}{\omega_1}\right) + 2\nu\pi i \tag{2}$$

gibt der Ansatz

$$z_\nu = \nu\,\omega_2 + z^*,$$

daß z^* mit wachsendem ν gegen Null strebt. Es ist linker Hand in (2), also aus den Anfangsgliedern der Potenzreihe eine Größenschätzung zu gewinnen; für konstante α und kleine z^* gilt

$$\left\{\begin{aligned} 2\nu\pi i &= \frac{\omega_1}{z^*} + \alpha z + O(z^2), \\ z^* &= -\frac{\omega_1 i}{2\nu\pi} + O(\nu^{-2}), \\ z_\nu &= \nu\,\omega_2 - \frac{\omega_1 i}{2\nu\pi} + O(\nu^{-2}). \end{aligned}\right. \tag{3}$$

Eine schärfere Schätzung zeigt, daß nicht nur für große, sondern für alle natürliche μ, ν

$$z_\mu + z_\nu \not\equiv 0\,(\omega_2),$$

so daß für unsere Eigenwerte z_ν die Bedingung § 6 (4) sichergestellt[2]) ist. Nach § 6 (9) gilt dann für jedes $\varkappa + \lambda \neq 0$

$$\int_\alpha^{\alpha+\omega_1} du\, f_\varkappa f_\lambda = 0.$$

Zu bestimmen bleibt noch der in § 6 (5′) betrachtete Ausdruck

$$\int_\alpha^{\alpha+\omega_1} du\, f_\nu f_{-\nu} = \int_\alpha^{\alpha+\omega_1} du\, s\begin{pmatrix} x & y \\ u & \end{pmatrix} s\begin{pmatrix} -x, & -y \\ u & \end{pmatrix}, \text{ nach § 4 (1) und § 2 (10)} \tag{4}$$

$$= \int_\alpha^{\alpha+\omega_1} du[\wp(u) - \wp(z_\nu)] = -\omega_1\wp(z_\nu) - 2\eta$$

mit von ν unabhängigem η.

---

[2]) Anmerkung bei der Korrektur. Im Fall singulärer Periodenverhältnisse gestattet ein Transzendenzsatz von C. Siegel die Irrationalität von $\frac{J(z_\nu)}{J\left(\frac{\omega_2}{\omega_1}\right)}$ zu beweisen.

Um in üblicher Weise eine Normierung der orthogonalen Funktionenfolge f_ν zu vollziehen, löse man die Bestimmungsgleichung

$$-2\eta = \omega_1 \wp(\zeta) \tag{5}$$

nach ζ auf; dann ist

$$\wp(\zeta) - \wp(z_\nu) = f(\zeta, z_\nu)\, f(\zeta, -z_\nu),$$

und der Ansatz

$$\frac{f(u, z_\nu)}{f(\zeta, z_\nu)} = \psi_\nu \tag{6}$$

liefert

$$\frac{1}{\omega_1} \int_\alpha^{\alpha+\omega_1} du\, \psi_\nu \psi_{-\nu} = 1 \tag{7}$$

neben

$$\int_\alpha^{\alpha+\omega_1} du\, \psi_\varkappa \psi_\lambda = 0.$$

In § 2 (7′) wurde

$$\int_\alpha^{\alpha+\omega_1} du\, f(u, z) = \frac{2\pi i}{1 - e^{-\frac{2\pi i z}{\omega_1}}}$$

gefunden, was zu

$$\frac{1}{\omega_1} \int_\alpha^{\alpha+\omega_1} du\, \psi_\nu = \frac{2\pi i}{\omega_1 \left(1 - e^{-\frac{2\pi i}{\omega_1} z_\nu}\right) f(\zeta, z_\nu)} \tag{8}$$

führt und die Größenschätzung der Koeffizienten b_ν in Reihenentwicklungen des Typs

$$1 = \sum_{-\infty}^{+\infty} b_\nu \psi_\nu \tag{9}$$

ermöglicht.

Der formale Ansatz (9) gibt wegen (7), (8) nämlich

$$b_{-\nu} = \frac{1}{\omega_1} \int_\alpha^{\alpha+\omega_1} du\, \psi_\nu = \frac{2\pi i}{\omega_1 \left(1 - e^{-\frac{2\pi i}{\omega_1} z_\nu}\right) f(\zeta, z_\nu)},$$

$$b_\nu \cdot \psi_\nu = \frac{2\pi i f(u, z_\nu)}{\omega_1 \left(-1 + e^{\frac{2\pi i}{\omega_1} z_\nu}\right) \left[\frac{2\eta}{\omega_1} + \wp(z_\nu)\right]} \tag{10}$$

$$= \frac{\vartheta\left(\frac{u + z_\nu}{\omega_1}\right)}{\vartheta\left(\frac{u}{\omega_1}\right) \left\{-1 + e^{\frac{2\pi i}{\omega_1} z_\nu}\right\} \left[\frac{\omega_1 \eta}{\pi i} + \frac{\omega_1}{2\pi i} \wp(z_\nu)\right] \vartheta\left(\frac{z_\nu}{\omega_1}\right)}.$$

Nach § 4 (6′) gilt

$$\vartheta\left(\frac{z + \nu\omega_2}{\omega_1}\right) = e^{\frac{\pi i}{\omega_1}(2\nu z + \nu^2 \omega_1)} \vartheta\left(\frac{z}{\omega_1}\right),$$

und der z_ν enthaltende ϑ-Quotient wächst wegen (3) für große ν wie die Vergleichsfunktion

$$e^{\frac{2\pi i}{\omega_1}\nu u};$$

in (10) eingetragen

$$(11)\qquad \vartheta\left(\frac{u}{\omega_1}\right)\cdot b_\nu\cdot\psi_\nu = \frac{e^{\frac{2\pi i}{\omega_1}\nu u}\cdot 2\pi i\omega_1\cdot\nu}{\left\{-1+e^{2\pi i\nu\frac{\omega_2}{\omega_1}}\right\}\omega_2^2\cdot\nu^2}\left[1+O\left(\frac{1}{\nu}\right)\right]$$

$$= \frac{2\pi i\omega_1}{\omega_2^2}\,\frac{1}{\nu}\,\frac{1}{e^{\frac{2\pi i\nu}{\omega_1}(\omega_2-u)}-e^{-\frac{2\pi i\nu}{\omega_1}u}}\left[1+O\left(\frac{1}{\nu}\right)\right].$$

Die Exponentialgrößen im Nenner von (11) sichern die absolute Konvergenz von (9) im Innern des Streifens

$$(12)\qquad 0<|J(u)|<\frac{\omega_2}{i}.$$

Aber nach (10) konvergiert dieselbe Entwicklung auch noch bedingt auf dem Randstück

$$0<u<\omega_1.$$

Allgemeiner kann für ganze Funktionen $g(u)$, die in ihrem Wachstum eingeschränkt sind, der Konvergenzbeweis wie oben durchgeführt werden; für

$$0<\Re(u)<\omega_1;\qquad 0<|J(u)|<\frac{\omega_2}{i}$$

ist z. B. jede solche Transzendente des Geschlechts 0 oder 1 nach ψ-Funktionen entwickelbar

$$g(u)=\sum_{-\infty}^{+\infty}{}'\,\psi_\nu(u)\,\frac{1}{\omega_2}\int_\alpha^{\alpha+\omega_1} dv\,\psi_{-\nu}(v)\,g(v).$$

Ulm a. d. D., Juli 1930.

(Eingegangen am 2. August 1930.)

Zur Verallgemeinerung des Bertrandschen Postulates, daß zwischen x und $2x$ stets Primzahlen liegen.

Von

Robert Breusch in Freiburg i. Br.

Einleitung.

Aus dem Primzahlsatz

$$\lim_{x \to \infty} \frac{\pi(x) \log x}{x} = 1$$

folgt, daß für positives ε und genügend großes a zwischen x und $(1+\varepsilon) \cdot x$ stets Primzahlen liegen, falls $x \geqq a$ ist. Damit hat man aber noch keine Methode, zu einer gegebenen festen Zahl ε ein hinreichend großes a wirklich zahlenmäßig festzustellen.

Das wichtigste Ergebnis in dieser Richtung ist bisher das folgende[1]): Wenn $\varepsilon > \frac{1}{5}$ ist, so liegen für hinreichend großes a und $x \geqq a$ zwischen x und $(1+\varepsilon) \cdot x$ stets Primzahlen, und die hinreichend großen Zahlen a können hier verhältnismäßig einfach bestimmt werden. Für $\varepsilon = \frac{1}{4}$ z. B. hat Herr Prof. I. Schur die Berechnung durchgeführt[2]) und die Gültigkeit des Satzes für alle Zahlen $x \geqq 24$ nachgewiesen.

In Teil I der vorliegenden Arbeit soll nun der Nachweis geführt werden, daß für $x \geqq 48$ zwischen x und $\frac{9}{8} x$ stets Primzahlen liegen, und in Teil II soll gezeigt werden, daß für $x \geqq 7$ zwischen x und $2x$ stets Primzahlen einer jeden der vier Progressionen $3n+1$, $3n+2$, $4n+1$, $4n+3$ liegen.

Wesentlich werden bei den Beweisen genaue zahlenmäßige Angaben über Anzahl und Lage der Nullstellen der Riemannschen ζ-Funktion und einiger Dirichletscher L-Reihen benützt, und zwar im einzelnen: von Backlund eine Abschätzung der Anzahl $N(T)$ der nicht-trivialen Null-

[1]) Vgl. Landau, Handbuch der Lehre von der Verteilung der Primzahlen, Leipzig 1909, **1**, § 22.

[2]) I. Schur, Einige Sätze über Primzahlen mit Anwendungen auf Irreduzibilitätsfragen I, Sitzungsberichte der preuß. Akad. d. Wissensch., phys.-math. Klasse 1929, S. 128.

stellen von $\zeta(s)$ mit Ordinate zwischen 0 und T;[3]) von demselben Verfasser ein Beweis, daß die Nullstellen von $\zeta(s)$ mit Ordinate zwischen 0 und 200 die Abszisse $\frac{1}{2}$ haben, sowie genaue Angaben über die Lage dieser Nullstellen;[4]) von J. Großmann Angaben über die ersten Nullstellen von $\zeta(s)$ und einigen L-Reihen.[5]) Wo für zahlenmäßige Durchführung der Berechnungen wirklich Kenntnis spezieller Primzahlen erforderlich war, wurden die Tabellen von D. N. Lehmer[6]) benützt.

Besonderen Dank schulde ich Herrn Prof. I. Schur, der die Anregung zu der Arbeit gab, und Herrn Prof. Loewy, der ihr großes Interesse entgegenbrachte.

Erster Teil.

§ 1.

Es sei in üblicher Bezeichnungsweise

$$\vartheta(x) = \sum_{p \leqq x} \log p \text{ und } \psi(x) = \sum_{p^m \leqq x} \log p \quad (m \text{ positiv ganzzahlig});$$

p soll dabei, wie immer im folgenden, Primzahlwerte annehmen. Dann ist

$$\psi(x) = \vartheta(x) + \vartheta(\sqrt{x}) + \vartheta(\sqrt[3]{x}) + \ldots,$$
$$\psi(x) - 2\psi(\sqrt{x}) = \vartheta(x) - \vartheta(\sqrt{x}) + \vartheta(\sqrt[3]{x}) - \ldots.$$

Daraus folgt

$$\psi(x) - 2\psi(\sqrt{x}) \leqq \vartheta(x) \leqq \psi(x). \tag{1}$$

Setzt man ferner $s = \sigma + it$ und

$$a_n = \psi(n) - \psi(n-1), \tag{2}$$

dann wird für $\sigma > 1$

$$\frac{\zeta'(s)}{\zeta(s)} = -\sum_{\nu=1}^{\infty} \frac{a_\nu}{\nu^s}.\ ^7) \tag{3}$$

Dabei ist $\zeta(s)$ die Riemannsche ζ-Funktion, die für $\sigma > 1$ definiert ist durch die Reihe $\zeta(s) = \sum_{\nu=1}^{\infty} \frac{1}{\nu^s}$.

---

[3]) Backlund, Über die Nullstellen der Riemannschen ζ-Funktion, Acta mathematica **41** (1918), S. 355.

[4]) Backlund, Über die Nullstellen der Riemannschen ζ-Funktion, Dissertation Helsingfors 1916.

[5]) J. Großmann, Über die Nullstellen der Riemannschen ζ-Funktion und der Dirichletschen L-Funktionen, Dissertation Göttingen 1913.

[6]) D. N. Lehmer, List of prime numbers from 1 to 10006721, Washington 1914.

[7]) Vgl. Landau, Handbuch (a. a. O.) **1**, § 33, S. 128. An Stelle von a_n log n steht hier a_n.

Weiter gilt der Satz: Wenn unter $\int_{(2)}$ das in der Ebene der komplexen Zahlen längs der ganzen Geraden $\sigma = 2$ erstreckte Integral verstanden wird, dann ist[8])

$$\frac{1}{2\pi i}\int_{(2)}\frac{y^s}{s^4}\,ds = \begin{cases} 0 & \text{für } 0 < y < 1, \\ \frac{1}{6}\log^3 y & \text{für } y \geqq 1. \end{cases} \tag{4}$$

Also ist

$$-\frac{1}{2\pi i}\int_{(2)}\frac{x^s}{s^4}\frac{\zeta'(s)}{\zeta(s)}\,ds = \frac{1}{2\pi i}\int_{(2)}\frac{x^s}{s^4}\sum_{\nu=1}^{\infty}\frac{a_\nu}{\nu^s}\,ds = \sum_{\nu=1}^{\infty}a_\nu\frac{1}{2\pi i}\int_{(2)}\frac{\left(\frac{x}{\nu}\right)^s}{s^4}\,ds$$

oder nach (4)

$$-\frac{1}{2\pi i}\int_{(2)}\frac{x^s}{s^4}\frac{\zeta'(s)}{\zeta(s)}\,ds = \frac{1}{6}\sum_{\nu\leqq x}a_\nu\log^3\frac{x}{\nu}. \tag{5}$$

§ 2.

Es sei für das folgende abgekürzt

$$\alpha(x) = \sum_{\nu\leqq x}a_\nu\log^3\frac{x}{\nu}.$$

Dann lautet Formel (5) in der neuen Bezeichnungsweise

$$\frac{1}{6}\alpha(x) = -\frac{1}{2\pi i}\int_{(2)}\frac{x^s}{s^4}\frac{\zeta'(s)}{\zeta(s)}ds. \tag{5'}$$

Der Integrand soll zunächst in der Ebene der komplexen Zahlen integriert werden über den Rand des Rechtecks mit den Ecken

$$2 - iT, \quad 2 + iT, \quad -(2m+1) + iT, \quad -(2m+1) - iT.$$

Dabei soll m positiv ganzzahlig sein und T positiv so gewählt werden, daß keine Nullstelle von $\zeta(s)$ die Ordinate T hat. Man erhält dann

$$\frac{1}{2\pi i}\int_{2-iT}^{2+iT}\frac{x^s}{s^4}\frac{\zeta'(s)}{\zeta(s)}\,ds = \frac{1}{2\pi i}\int_{-(2m+1)-iT}^{-(2m+1)+iT}\frac{x^s}{s^4}\frac{\zeta'(s)}{\zeta(s)}\,ds + \frac{1}{2\pi i}\int_{-(2m+1)+iT}^{2+iT}\frac{x^s}{s^4}\frac{\zeta'(s)}{\zeta(s)}\,ds$$
$$-\frac{1}{2\pi i}\int_{-(2m+1)-iT}^{2-iT}\frac{x^s}{s^4}\frac{\zeta'(s)}{\zeta(s)}ds + \sum \text{Resid.},$$

wobei die Summe erstreckt werden soll über alle Pole des Integranden im Rechteck.

---

[8]) Vgl. Landau, Vorlesungen über Zahlentheorie, Leipzig 1927, 2, S. 21, Satz 377; dort ist mit s^2 statt s^4 gerechnet, doch läßt sich der Beweis im vorliegenden Fall ganz analog durchführen.

Die Absolutwerte der beiden Integrale $\int_{-(2m+1)+iT}^{2+iT}$ und $\int_{-(2m+1)-iT}^{2-iT}$ sind kleiner als[9]) $\frac{cx^3 \log T}{(x-1)T^4}$, beide Integrale gehen also für $x>1$ mit wachsendem T gegen 0. c bedeute hierbei, ebenso wie nachher c_1, eine Konstante.

Weiter ist[10]) für $\sigma = -(2m+1)$

$$\left|\frac{\zeta'(s)}{\zeta(s)}\right| < c_1 \log|s| < c_1 \log(|t|+2m+1),$$

also

$$\left|\int_{-(2m+1)-i\infty}^{-(2m+1)+i\infty} \frac{x^s}{s^4}\frac{\zeta'(s)}{\zeta(s)}\,ds\right| < c_1 x^{-(2m+1)} \int_{-\infty}^{+\infty} \frac{\log(|t|+2m+1)}{(2m+1)^4+t^4}\,dt.$$

Für $x>1$ geht das Integral mit wachsendem m gegen 0.

Daraus folgt

$$\frac{1}{2\pi i}\int_{(2)} \frac{x^s}{s^4}\frac{\zeta'(s)}{\zeta(s)}\,ds = \sum \text{Resid.},$$

wobei die Summe jetzt erstreckt werden soll über alle Pole des Integranden in der Halbebene $\sigma<2$.

Voraussetzung für die letzte Formel ist, daß $\sum$ Resid. überhaupt konvergiert. Die Konvergenz dieser Summe soll zunächst gezeigt werden. Der Integrand $\frac{x^s}{s^4}\frac{\zeta'(s)}{\zeta(s)}$ hat Pole an folgenden Stellen:

a) $s=1$, Residuum $-x$.[11])

b) $s=\varrho$, wo ϱ eine der nicht-trivialen Nullstellen von $\zeta(s)$ ist, d. h. eine Nullstelle mit $0<\sigma<1$. Das Residuum von $\frac{\zeta'(s)}{\zeta(s)}$ an einer solchen Stelle ist k, wenn es sich um eine k-fache Nullstelle handelt. Also ist die Summe der zugehörigen Residuen gleich $\sum_{\varrho}\frac{x^\varrho}{\varrho^4}$; dabei sind k-fache Wurzeln k-mal zu nehmen. Die Summe ist absolut konvergent.

c) $s=-2n$, $n=1,2,3,\ldots$; Residuum von $\frac{\zeta'(s)}{\zeta(s)}$ an einer solchen Stelle ist 1, also ist die Summe der Residuen gleich $\sum_{n=1}^{\infty}\frac{x^{-2n}}{(2n)^4}$. Auch diese Summe konvergiert, falls $x \geqq 1$ ist.

d) $s=0$. $\frac{\zeta'(s)}{\zeta(s)}$ ist hier regulär

$$\frac{\zeta'(s)}{\zeta(s)} = a_0 + a_1 s + a_2 s^2 + \ldots, \qquad x^s = 1 + s\log x + \frac{s^2\log^2 x}{2!} + \ldots;$$

---

[9]) Vgl. Landau, Vorl. ü. Zahlentheorie 2, S. 119—120.

[10]) Ebenda, S. 111, Satz 448.

[11]) Landau, Handbuch 1, § 48, S. 179.

also

$$\frac{x^s}{s^4}\frac{\zeta'(s)}{\zeta(s)} = \ldots + \frac{1}{s}(b_0 + b_1 \log x + b_2 \log^2 x + b_3 \log^3 x) + \ldots .$$

Das Residuum an der Stelle 0 ist also

$$b_0 + b_1 \log x + b_2 \log^2 x + b_3 \log^3 x.$$

Der Wert der Konstanten b_i wird sich als unwesentlich herausstellen. Aus (5′) folgt nun

$$(6) \quad \frac{1}{6}\alpha(x) = x - \sum_{\varrho}\frac{x^{\varrho}}{\varrho^4} - \sum_{n=1}^{\infty}\frac{x^{-2n}}{(2n)^4} - b_0 - b_1 \log x - b_2 \log^2 x - b_3 \log^3 x.$$

§ 3.

y sei eine später zu bestimmende positive Größe. Es soll folgender Ausdruck gebildet werden:

$$\Delta = \alpha((1+y)^4 x) - 4\alpha((1+y)^3 x) + 6\alpha((1+y)^2 x) - 4\alpha((1+y)x) + \alpha(x)$$

oder:

$$\Delta = \sum_{\nu \leqq (1+y)^4 x} a_\nu \log^3 \frac{(1+y)^4 x}{\nu} - 4 \sum_{\nu \leqq (1+y)^3 x} a_\nu \log^3 \frac{(1+y)^3 x}{\nu} + 6 \sum_{\nu \leqq (1+y)^2 x} a_\nu \log^3 \frac{(1+y)^2 x}{\nu}$$
$$- 4 \sum_{\nu \leqq (1+y)x} a_\nu \log^3 \frac{(1+y)x}{\nu} + \sum_{\nu \leqq x} a_\nu \log^3 \frac{x}{\nu}.$$

Ordnet man nach Größe von ν, so wird

$$\sum_{\nu \leqq x} a_\nu \left[\log^3 \frac{(1+y)^4 x}{\nu} - 4 \log^3 \frac{(1+y)^3 x}{\nu} + 6 \log^3 \frac{(1+y)^2 x}{\nu} - 4 \log^3 \frac{(1+y)x}{\nu} + \log^3 \frac{x}{\nu}\right]$$
$$+ \sum_{x < \nu \leqq (1+y)x} a_\nu \left[\log^3 \frac{(1+y)^4 x}{\nu} - 4 \log^3 \frac{(1+y)^3 x}{\nu} + 6 \log^3 \frac{(1+y)^2 x}{\nu} - 4 \log^3 \frac{(1+y)x}{\nu}\right]$$
$$+ \sum_{(1+y)x < \nu \leqq (1+y)^2 x} a_\nu \left[\log^3 \frac{(1+y)^4 x}{\nu} - 4 \log^3 \frac{(1+y)^3 x}{\nu} + 6 \log^3 \frac{(1+y)^2 x}{\nu}\right]$$
$$+ \sum_{(1+y)^2 x < \nu \leqq (1+y)^3 x} a_\nu \left[\log^3 \frac{(1+y)^4 x}{\nu} - 4 \log^3 \frac{(1+y)^3 x}{\nu}\right] + \sum_{(1+y)^3 x < \nu \leqq (1+y)^4 x} a_\nu \log^3 \frac{(1+y)^4 x}{\nu}.$$

Zunächst sollen die einzelnen Koeffizienten der a_ν berechnet oder abgeschätzt werden. Der erste Koeffizient hat den Wert 0; mit Rücksicht auf spätere Anwendungen soll hier gleich gezeigt werden: wenn μ positiv ganzzahlig $\leqq 4$ ist, dann ist

$$\Delta_\mu = \log^\mu \frac{(1+y)^4 x}{\nu} - 4 \log^\mu \frac{(1+y)^3 x}{\nu} + 6 \log^\mu \frac{(1+y)^2 x}{\nu} - 4 \log^\mu \frac{(1+y)x}{\nu} + \log^\mu \frac{x}{\nu}$$
$$= \begin{cases} 0 \text{ für } \mu \leqq 3, \\ 24 \log^4 (1+y) \text{ für } \mu = 4. \end{cases}$$

Beweis. Es ist

$$\Lambda_\mu = \left[\log(1+y)^4 + \log\frac{x}{\nu}\right]^\mu - 4\left[\log(1+y)^3 + \log\frac{x}{\nu}\right]^\mu + 6\left[\log(1+y)^2 + \log\frac{x}{\nu}\right]^\mu$$
$$- 4\left[\log(1+y) + \log\frac{x}{\nu}\right]^\mu + \left[\log\frac{x}{\nu}\right]^\mu$$
$$= \log^\mu\frac{x}{\nu}[1-4+6-4+1]$$
$$+ \binom{\mu}{1}\log^{\mu-1}\frac{x}{\nu}[\log(1+y)^4 - 4\log(1+y)^3 + 6\log(1+y)^2 - 4\log(1+y)]$$
$$+ \binom{\mu}{2}\log^{\mu-2}\frac{x}{\nu}[\log^2(1+y)^4 - 4\log^2(1+y)^3 + 6\log^2(1+y)^2 - 4\log^2(1-$$
$$+ \binom{\mu}{3}\log^{\mu-3}\frac{x}{\nu}[\log^3(1+y)^4 - 4\log^3(1+y)^3 + 6\log^3(1+y)^2 - 4\log^3(1-$$
$$+ \binom{\mu}{4}\log^{\mu-4}\frac{x}{\nu}[\log^4(1+y)^4 - 4\log^4(1+y)^3 + 6\log^4(1+y)^2 - 4\log^4(1-$$
$$= \log^\mu\frac{x}{\nu}[1-4+6-4+1] + \binom{\mu}{1}\log^{\mu-1}\frac{x}{\nu}\log(1+y)[4-4\cdot3+6\cdot2-4$$
$$+ \binom{\mu}{2}\log^{\mu-2}\frac{x}{\nu}\log^2(1+y)[16-4\cdot9+6\cdot4-4]$$
$$+ \binom{\mu}{3}\log^{\mu-3}\frac{x}{\nu}\log^3(1+y)[64-4\cdot27+6\cdot8-4]$$
$$+ \binom{\mu}{4}\log^{\mu-4}\frac{x}{\nu}\log^4(1+y)[256-4\cdot81+6\cdot16-4]$$
$$= \binom{\mu}{4}\log^{\mu-4}\frac{x}{\nu}\log^4(1+y)\cdot24 = \begin{cases} 0 \text{ für } \mu \leqq 3, \\ 24\log^4(1+y) \text{ für } \mu = 4. \end{cases}$$

Der zweite Koeffizient von a_ν (d. h. der in $\sum_{x<\nu\leqq(1+y)x}$) hat demnach den Wert $-\log^3\frac{x}{\nu}$; da ν hier zwischen x und $(1+y)x$ liegt, so ist der Koeffizient $\leqq \log^3(1+y)$.

Der dritte Koeffizient ist

$$\log^3\frac{(1+y)^4x}{\nu} - 4\log^3\frac{(1+y)^3x}{\nu} + 6\log^3\frac{(1+y)^2x}{\nu}$$
$$= \left[\log(1+y)^2 + \log\frac{(1+y)^2x}{\nu}\right]^3 - 4\left[\log(1+y) + \log\frac{(1+y)^2x}{\nu}\right]^3 + 6\left[\log\frac{(1+y)}{\nu}\right.$$
$$= \log^3\frac{(1+y)^2x}{\nu}[1-4+6] + 3\log^2\frac{(1+y)^2x}{\nu}[\log(1+y)^2 - 4\log(1+y)]$$
$$+ 3\log\frac{(1+y)^2x}{\nu}[\log^2(1+y)^2 - 4\log^2(1+y)] + [\log^3(1+y)^2 - 4\log^3(1-$$
$$= 3\log^3\frac{(1+y)^2x}{\nu} - 6\log(1+y)\log^2\frac{(1+y)^2x}{\nu} + 4\log^3(1+y).$$

Da ν hier zwischen $(1+y)x$ und $(1+y)^2x$ liegt, so ist der Ausdruck $\leqq 4\log^3(1+y)$. Der vierte Koeffizient ist wieder $\leqq 4\log^3(1+y)$, der fünfte schließlich $\leqq \log^3(1+y)$.

Also ist insgesamt

$$(8) \qquad \Delta \leqq 4\log^3(1+y) \sum_{x<\nu\leqq(1+y)^4 x} a_\nu \leqq 4y^3 \sum_{x<\nu\leqq(1+y)^4 x} a_\nu .$$

Andrerseits ist nach (6)

$$\begin{aligned}
&x[(1+y)^4-4(1+y)^3+6(1+y)^2-4(1+y)+1]\\
&-\sum_{\varrho}\frac{x^\varrho[(1+y)^{4\varrho}-4(1+y)^{3\varrho}+6(1+y)^{2\varrho}-4(1+y)^\varrho+1]}{\varrho^4}\\
&-\sum_{n=1}^{\infty}\frac{x^{-2n}[(1+y)^{-4\cdot 2n}-4(1+y)^{-3\cdot 2n}+6(1+y)^{-2\cdot 2n}-4(1+y)^{-2n}+1]}{(2n)^4}\\
&-b_0[1-4+6-4+1]\\
&-b_1[\log(1+y)^4 x-4\log(1+y)^3 x+6\log(1+y)^2 x-4\log(1+y)x+\log x]\\
&-b_2[\log^2(1+y)^4 x-4\log^2(1+y)^3 x+6\log^2(1+y)^2 x-4\log^2(1+y)x+\log^2 x]\\
&-b_3[\log^3(1+y)^4 x-4\log^3(1+y)^3 x+6\log^3(1+y)^2 x-4\log^3(1+y)x+\log^3 x].
\end{aligned}$$

Die Koeffizienten der b_i sind nach (7) sämtlich 0, also ist

$$\frac{\Delta}{6}=y^4 x-\sum_{\varrho}\frac{x^\varrho[(1+y)^\varrho-1]^4}{\varrho^4}-\sum_{n=1}^{\infty}\frac{x^{-2n}[(1+y)^{-2n}-1]^4}{(2n)^4}.$$

Daraus und aus (8) folgt

$$(9) \qquad \frac{4}{6}y^3\cdot\sum_{x<\nu\leqq(1+y)^4 x} a_\nu \geqq y^4 x-\sum_{\varrho}\frac{x^\varrho[(1+y)^\varrho-1]^4}{\varrho^4}-\sum_{n=1}^{\infty}\frac{x^{-2n}[(1+y)^{-2n}-1]^4}{(2n)^4}.$$

§ 4.

Nun müssen die rechts stehenden Summen abgeschätzt werden. Zunächst die zweite unter der Voraussetzung $x>1$: Es ist

$$(10) \qquad \sum_{n=1}^{\infty}\frac{x^{-2n}[(1+y)^{2n}-1]^4}{(2n)^4}<\frac{1}{16x^2}\sum_{n=1}^{\infty}\frac{1}{n^4}<\frac{1}{12x^2}.$$

Weiter sei im folgenden $\varrho=\varrho_1+i\varrho_2$. Dann ist

$$\sum_{\varrho}\frac{x^\varrho[(1+y)^\varrho-1]^4}{\varrho^4}=\sum_{|\varrho_2|\geqq 200}\frac{x^\varrho[(1+y)^\varrho-1]^4}{\varrho^4}+\sum_{|\varrho_2|<200}\frac{x^\varrho[(1+y)^\varrho-1]^4}{\varrho^4}.$$

Für $|\varrho_2|<200$ ist $\varrho_1=\frac{1}{2}$.[12]) Die Abszissen der übrigen Nullstellen sind alle $\leqq 1$, mindestens die Hälfte ist $\leqq\frac{1}{2}$, da die Nullstellen zu $\sigma=\frac{1}{2}$ symmetrisch liegen. Da sie auch zu $t=0$ symmetrisch liegen, so ist

$$(11) \qquad \left|\sum_{|\varrho_2|\geqq 200}\frac{x^\varrho[(1+y)^\varrho-1]^4}{\varrho^4}\right|<(2+y)^4\left[x\sum_{\varrho_2\geqq 200}\frac{1}{|\varrho|^4}+\sqrt{x}\sum_{\varrho_2\geqq 200}\frac{1}{|\varrho|^4}\right].$$

[12]) Vgl. die in der Einleitung zitierten Arbeiten von Backlund.

Es sei $N(T)$ die Anzahl der Nullstellen mit Ordinate zwischen 0 und T, k-fache k-fach gezählt; dann ist

$$\sum_{\varrho_2 \geqq 200} \frac{1}{|\varrho|^4} \leqq \sum_{n=200}^{\infty} \frac{N(n+1)-N(n)}{n^4}. \tag{12}$$

Nun ist[12])

$$N(T) = \frac{T}{2\pi}\left(\log\frac{T}{2\pi} - 1\right) + \frac{7}{8} + Q(T), \tag{13}$$

oder abgekürzt:

$$N(T) = R(T) + Q(T).$$

Dabei gilt für $Q(T)$ [12])

$$|Q(T)| \leqq 0{,}137 \log T + 0{,}443 \log\log T + 4{,}350. \tag{14}$$

Also

$$\sum_{n=200}^{\infty} \frac{N(n+1)-N(n)}{n^4} = \sum_{n=200}^{\infty} \frac{R(n+1)-R(n)}{n^4} + \sum_{n=200}^{\infty} \frac{Q(n+1)-Q(n)}{n^4}.$$

Die zweite Summe wird weiter zerlegt:

$$\sum_{n=200}^{\infty} \frac{Q(n+1)-Q(n)}{n^4} = \sum_{n=200}^{999} \frac{Q(n+1)-Q(n)}{n^4} + \sum_{\varkappa=1}^{\infty} \sum_{n=\varkappa\cdot 1000}^{\varkappa\cdot 1000+999} \frac{Q(n+1)-Q(n)}{n^4}.$$

$\sum_{n=200}^{999} \frac{Q(n+1)-Q(n)}{n^4}$ läßt sich etwa auf folgende Weise abschätzen: $Q(n+1)-Q(n)$ kann positiv oder negativ sein; jedenfalls kann man sagen:

$$Q(1000) - Q(999) \leqq M_1,$$

wobei M_1 die größere der zwei Zahlen 0 und $Q(1000) - Q(999)$ bedeuten soll. Also ist

$$\frac{Q(999)-Q(998)}{998^4} + \frac{Q(1000)-Q(999)}{999^4} \leqq \frac{Q(999)-Q(998)+M_1}{998^4} \leqq \frac{M_2}{998^4},$$

wobei nun M_2 die größere der zwei Zahlen 0 und $Q(999) - Q(998) + M_1$ sein soll. Entsprechend weiter

$$\sum_{n=997}^{999} \frac{Q(n+1)-Q(n)}{n^4} \leqq \frac{M_3}{997^4},$$

wobei M_3 die größere der zwei Zahlen 0 und $Q(998) - Q(997) + M_2$ ist, usw. Schließlich wird

$$\sum_{n=200}^{999} \frac{Q(n+1)-Q(n)}{n^4} \leqq \frac{M_{800}}{200^4}.$$

M_{800} ist dabei die größte der Zahlen

$$Q(1000) - Q(200);\ Q(999) - Q(200);\ \ldots;\ Q(201) - Q(200);\ 0.$$

Nach (14) ist also sicher

$$\sum_{n=200}^{999} \frac{Q(n+1)-Q(n)}{n^4} \leqq \frac{1}{200^4} \cdot [0{,}137\,(\log 200 + \log 1000) + 0{,}443\,(\log\log 200 + \log\log 1000) + 8{,}700] < \frac{12}{200^4}.$$

Entsprechend gilt

$$\sum_{1000}^{\varkappa+999} \frac{Q(n+1)-Q(n)}{n^4} \leqq \frac{1}{\varkappa^4 \cdot 10^{12}} \cdot [0{,}274 \log(\varkappa+1)\,1000 + 0{,}886 \log\log(\varkappa+1)\,1000 + 8{,}700] < \frac{1}{\varkappa^4 \cdot 10^{12}} \cdot [\log(\varkappa+1) + 14] < \frac{1}{\varkappa^4 \cdot 10^{12}} \cdot [\log \varkappa + 15].$$

Also ist

$$\sum_{n=200}^{\infty} \frac{Q(n+1)-Q(n)}{n^4} < \frac{12}{200^4} + \frac{1}{10^{12}} \sum_{\varkappa=1}^{\infty} \frac{\log \varkappa + 15}{\varkappa^4} < \frac{12}{200^4} + \frac{21}{10^{12}},$$

d. h.

$$\sum_{n=200}^{\infty} \frac{Q(n+1)-Q(n)}{n^4} < 0{,}76 \cdot 10^{-8}. \tag{15}$$

Weiter ist nach (13)

$$R(n+1) - R(n) = \frac{n+1}{2\pi}\left[\log\frac{n+1}{2\pi} - 1\right] - \frac{n}{2\pi}\left[\log\frac{n}{2\pi} - 1\right] = \frac{1}{2\pi}\left[\log\frac{n}{2\pi} - 1\right] + \frac{n+1}{2\pi}\log\left(1+\frac{1}{n}\right) < \frac{1}{2\pi}\log\frac{n}{2\pi} + \frac{1}{2\pi n}.$$

Also ist

$$\sum_{n=200}^{\infty} \frac{R(n+1)-R(n)}{n^4} < \frac{1}{2\pi} \sum_{n=200}^{\infty} \frac{\log\frac{n}{2\pi} + \frac{1}{n}}{n^4} < 2{,}56 \cdot 10^{-8}.$$

Daraus, aus (12) und aus (15) folgt

$$\sum_{\varrho_2 \geqq 200} \frac{1}{|\varrho|^4} < (2{,}56 + 0{,}76) \cdot 10^{-8} = 3{,}32 \cdot 10^{-8}. \tag{16}$$

Also ist nach (9)

$$\frac{2}{3} y^3 \sum_{x < \nu \leqq (1+y)^4 x} a_\nu > x \cdot [y^4 - (2+y)^4 \cdot 3{,}32 \cdot 10^{-8}] - \sqrt{x} \cdot \left[(2+y)^4 \cdot 3{,}32 \cdot 10^{-8} + 2 \cdot \sum_{0 < \varrho_2 < 200} \frac{|(1+y)^\varrho - 1|^4}{|\varrho|^4}\right] - \frac{1}{12x^2}. \tag{17}$$

Um auch diese letzte Summe noch abzuschätzen, braucht man genauere Angaben über die Lage der Nullstellen mit Ordinate zwischen 0 und 200.

Es ist[13])

$$(18)\quad 14 \leqq \varrho_2^{(1)} \leqq 20 \leqq \varrho_2^{(2)} \leqq 24 \leqq \varrho_2^{(3)} \leqq 30 \leqq \varrho_2^{(4)} \text{ und } \varrho_2^{(5)} \leqq 34 \leqq \varrho_2^{(6)} \text{ bis } \varrho_2^{(10)} \leqq 50 \leqq \varrho_2^{(11)} \text{ bis } \varrho_2^{(29)} \leqq 100 \leqq \varrho_2^{(30)} \text{ bis } \varrho_2^{(79)} \leqq 200\,.$$

Es sei im folgenden $y = \frac{1}{34}$; dann ist $(1+y)^4 < \frac{9}{8}$, $(2+y)^4 < 17$. Weiter ist dann für die Nullstellen mit $|\varrho_2| < 200$

$$(1+y)^\varrho - 1 = \left(\frac{35}{34}\right)^{\frac{1}{2}} \cos\left(\varrho_2 \log\frac{35}{34}\right) - 1 + i\left(\frac{35}{34}\right)^{\frac{1}{2}} \sin\left(\varrho_2 \log\frac{35}{34}\right).$$

Sobald der Realteil negativ ist, nimmt sein Absolutbetrag zu mit wachsendem ϱ_2, der des Imaginärteils nimmt ebenfalls zu mit wachsendem ϱ_2, und zwar so lange, als $\varrho_2 \cdot \log\frac{35}{34} < \frac{\pi}{2}$ bleibt, also sicher für $\varrho_2 \leqq 50$. Daher ist nach (18)

$$\begin{aligned}\sum_{0<\varrho_2<200} \frac{|(1+y)^\varrho - 1|^4}{|\varrho|^4} &< \left|\left(\frac{35}{34}\right)^{\frac{1}{2}} \cos\left(34 \log\frac{35}{34}\right) - 1 + i\left(\frac{35}{34}\right)^{\frac{1}{2}} \sin\left(34 \log\frac{35}{34}\right)\right|^4 \\ &\quad \times \left[\frac{1}{14^4} + \frac{1}{20^4} + \frac{1}{24^4} + \frac{2}{30^4}\right] \\ &\quad + \left|\left(\frac{35}{34}\right)^{\frac{1}{2}} \cos\left(50 \log\frac{35}{34}\right) - 1 + i\left(\frac{35}{34}\right)^{\frac{1}{2}} \sin\left(50 \log\frac{35}{34}\right)\right|^4 \cdot \frac{5}{34^4} \\ &\quad + \left[1 + \left(\frac{35}{34}\right)^{\frac{1}{2}}\right]^4 \cdot \left[\frac{19}{50^4} + \frac{50}{100^4}\right] \\ &< 1 \cdot \left[\frac{1}{14^4} + \frac{1}{20^4} + \frac{1}{24^4} + \frac{2}{30^4}\right] + 4\,\frac{5}{34^4} + 16{,}5 \cdot \left[\frac{19}{50^4} + \frac{50}{100^4}\right]\end{aligned}$$

oder

$$(19)\qquad \sum_{0<\varrho_2<200} \frac{|(1+y)^\varrho - 1|^4}{|\varrho|^4} < 11{,}12 \cdot 10^{-5}.$$

Aus (17) und (19) folgt für $y = \frac{1}{34}$

$$\begin{aligned}\frac{2}{3} \cdot \frac{1}{34^3} \sum_{x<\nu \leqq \frac{9}{8}x} a_\nu &> x\left(\frac{1}{34^4} - 17 \cdot 3{,}32 \cdot 10^{-8}\right) - \sqrt{x}\,(17 \cdot 3{,}32 \cdot 10^{-8} + 2 \cdot 11{,}12 \cdot 10^{-5}) - \\ &> 18{,}38 \cdot 10^{-8} x - 22{,}3 \cdot 10^{-5} \cdot \sqrt{x} - \frac{1}{12x^2}.\end{aligned}$$

Es sei $x > 4000000$; dann ist

$$\frac{2}{3} \cdot \frac{1}{34^3} \sum_{x<\nu \leqq \frac{9}{8}x} a_\nu > (18{,}38 - 11{,}16) \cdot 10^{-8} x > 7 \cdot 10^{-8} x\,,$$

also

$$(20)\qquad \sum_{x<\nu \leqq \frac{9}{8}x} a_\nu > 7 \cdot \frac{3}{2} \cdot 34^3 \cdot 10^{-8} x > 40 \cdot 10^{-4} x\,.$$

---

[13]) Vgl. die schon zitierten Arbeiten von Backlund und Großmann.

Setzt man hier für a_ν seinen Wert nach (2) ein, dann erhält man

$$\psi\left(\tfrac{9}{8}x\right) - \psi(x) > 40 \cdot 10^{-4} x$$

oder wegen (1)

$$\vartheta\left(\tfrac{9}{8}x\right) - \vartheta(x) > \psi\left(\tfrac{9}{8}x\right) - \psi(x) - 2\psi\left(\sqrt{\tfrac{9}{8}x}\right) > 40 \cdot 10^{-4} x - 2\psi\left(\sqrt{\tfrac{9}{8}x}\right).$$

Nun ist[14])

$$\psi(z) < 1{,}106\, z + 3 \log^2 z + 8 \log z + 5.$$

Also ist für $x \geqq 4000000$

$$\psi\left(\sqrt{\tfrac{9}{8}x}\right) < 1{,}25 \cdot \sqrt{\tfrac{9}{8}x} < 1{,}34 \cdot \sqrt{x} < 6{,}7 \cdot 10^{-4} x,$$

d. h.

$$\vartheta\left(\tfrac{9}{8}x\right) - \vartheta(x) > (40 - 13{,}4) \cdot 10^{-4} x > 26 \cdot 10^{-4} x > 0. \tag{21}$$

Zwischen x und $\frac{9}{8}x$ liegen also für $x \geqq 4 \cdot 10^6$ stets Primzahlen. Mit Hilfe der Lehmerschen Tabellen wird die Gültigkeit des Satzes festgestellt für $x \geqq 48$. Man überzeugt sich nämlich leicht, daß in den bis 10^7 reichenden Tabellen die Differenz zweier aufeinanderfolgender Primzahlen stets kleiner als 1000 ist, so daß der Satz für $8000 < x < 4 \cdot 10^6$ offenbar gilt. Entsprechend sieht man weiter, daß der Satz richtig ist für $800 < x \leqq 8000$, weil für Werte unterhalb 8000 die Differenz zweier aufeinanderfolgender Primzahlen < 100 ist. Unterhalb 800 wieder ist diese Differenz < 20, der Satz gilt also für $160 < x \leqq 800$. Und schließlich ist von den im folgenden angeführten Primzahlen jede kleiner als das $\frac{9}{8}$-fache der vorhergehenden, die erste ist kleiner als $\frac{9}{8} \cdot 48$ und die letzte größer als 160:

53, 59, 61, 67, 73, 79, 83, 89, 97, 109, 113, 127, 139, 151, 167.

Damit ist gezeigt: für $x \geq 48$ liegen zwischen x und $\frac{9}{8}x$ stets Primzahlen.

Zweiter Teil.

§ 1.

χ sei irgendein Charakter mod k, k eine beliebige natürliche Zahl. $L(s, \chi)$ sei die zugehörige Dirichletsche L-Funktion, die für $\sigma > 1$ definiert ist durch

$$L(s, \chi) = \sum_{\nu=1}^{\infty} \frac{\chi(\nu)}{\nu^s}.$$

[14]) Landau, Handbuch 1, § 22, S. 90.

Dann ist[15])

$$\frac{L'(s,\chi)}{L(s,\chi)} = -\sum_{\nu=1}^{\infty} \frac{\chi(\nu) a_\nu}{\nu^s}$$

für $\sigma > 1$ absolut konvergent. a_ν ist dabei durch die Relation (2) im ersten Teil erklärt. Also gilt

$$(1) \quad -\frac{1}{2\pi i}\int_{(2)} \frac{x^s}{s^4}\frac{L'(s,\chi)}{L(s,\chi)}\,ds = \frac{1}{2\pi i}\int_{(2)} \frac{x^s}{s^4}\sum_{\nu=2}^{\infty}\frac{\chi(\nu)\cdot a_\nu}{\nu^s}\,ds$$

$$= \sum_{\nu=2}^{\infty}\chi(\nu)\,a_\nu \frac{1}{2\pi i}\int_{(2)}\frac{\left(\frac{x}{\nu}\right)^s}{s^4}\,ds = \frac{1}{6}\sum_{\nu\leqq x}\chi(\nu)\,a_\nu \log^3\frac{x}{\nu},$$

ganz analog zu den Entwicklungen in § 1 des ersten Teiles.

Es sei $h = \varphi(k) =$ Anzahl der zu k teilerfremden Zahlen zwischen 0 und k. Dann ist für zu k teilerfremdes l

$$\sum_{\chi}\frac{\chi(n)}{\chi(l)} = \begin{cases} 0 & \text{für } n \not\equiv l \pmod k, \\ h & \text{für } n \equiv l \pmod k. \end{cases}$$

Hierbei soll die Summe erstreckt werden über alle Charaktere mod k. Sei nämlich r so beschaffen, daß $r\cdot l \equiv 1 \pmod k$ ist. Dann wird $\chi(r)\cdot\chi(l) = 1$ und[16])

$$\sum_{\chi}\frac{\chi(n)}{\chi(l)} = \sum_{\chi}\chi(r)\cdot\chi(n) = \sum_{\chi}\chi(r\cdot n) = \begin{cases} 0 & \text{für } r\cdot n \not\equiv 1 \pmod k, \\ h & \text{für } r\cdot n \equiv 1 \pmod k, \end{cases}$$

$$= \begin{cases} 0 & \text{für } n \not\equiv l \pmod k, \\ h & \text{für } n \equiv l \pmod k. \end{cases}$$

Daraus und aus (1) folgt

$$-\sum_{\chi}\frac{1}{\chi(l)}\cdot\frac{1}{2\pi i}\int_{(2)}\frac{x^s}{s^4}\frac{L'(s,\chi)}{L(s,\chi)}\,ds = \frac{1}{6}\sum_{\nu\leqq x} a_\nu \log^3\frac{x}{\nu}\sum_{\chi}\frac{\chi(\nu)}{\chi(l)}$$

oder

$$(2) \quad -\sum_{\chi}\frac{1}{\chi(l)}\cdot\frac{1}{2\pi i}\int_{(2)}\frac{x^s}{s^4}\frac{L'(s,\chi)}{L(s,\chi)}\,ds = \frac{h}{6}\sum_{\substack{\nu\leqq x \\ \nu\equiv l \pmod k}} a_\nu \log^3\frac{x}{\nu}.$$

§ 2.

Es wird sich auch hier als notwendig erweisen, die Anzahl $N(T)$ der Nullstellen von $L(s,\chi)$ mit Ordinate zwischen 0 und T abzuschätzen. Für unsere Zwecke genügt eine etwas rohere Abschätzung, die ganz analog zu der in Landaus Vorlesungen über Zahlentheorie 2, S. 87 hergeleitet wird.

---

[15]) Vgl. Landau, Handbuch 1, S. 420.

[16]) Ebenda, S. 408, Satz 4.

Es sei χ speziell ein reeller, eigentlicher Charakter, für den $\chi(-1) = -1$ ist. Dann ist[17])

$$L(s,\chi) = a e^{bs} \frac{1}{\Gamma\left(\frac{s+1}{2}\right)} \prod_{\varrho^*} \left(1 - \frac{s}{\varrho^*}\right) e^{\frac{s}{\varrho^*}}.$$

a und b sind dabei Konstante, und ϱ^* durchlaufe die nicht-trivialen Nullstellen von $L(s,\chi)$. Daraus folgt

$$\frac{L'(s,\chi)}{L(s,\chi)} = b - \frac{1}{2}\frac{\Gamma'\left(\frac{s+1}{2}\right)}{\Gamma\left(\frac{s+1}{2}\right)} + \sum_{\varrho^*}\left(\frac{1}{s-\varrho^*} + \frac{1}{\varrho^*}\right). \tag{3}$$

Für $s = 0$ erhält man hieraus

$$b = \frac{L'(0,\chi)}{L(0,\chi)} + \frac{1}{2}\frac{\Gamma'\left(\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)}.$$

Weiter sei

$$\xi(s,\chi) = \left(\frac{\pi}{k}\right)^{-\frac{s+1}{2}} \Gamma\left(\frac{s+1}{2}\right) L(s,\chi).$$

Dann ist[18]), da χ als reeller Charakter mit seinem konjugiert komplexen identisch ist,

$$\frac{\xi'(s,\chi)}{\xi(s,\chi)} = -\frac{\xi'(1-s,\chi)}{\xi(1-s,\chi)}.$$

Also

$$\frac{L'(s,\chi)}{L(s,\chi)} + \frac{1}{2}\frac{\Gamma'\left(\frac{s+1}{2}\right)}{\Gamma\left(\frac{s+1}{2}\right)} - \frac{1}{2}\log\frac{\pi}{k} = -\frac{L'(1-s,\chi)}{L(1-s,\chi)} - \frac{1}{2}\frac{\Gamma'\left(\frac{2-s}{2}\right)}{\Gamma\left(\frac{2-s}{2}\right)} + \frac{1}{2}\log\frac{\pi}{k},$$

$$b = \frac{L'(0,\chi)}{L(0,\chi)} + \frac{1}{2}\frac{\Gamma'\left(\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)} = -\frac{L'(1,\chi)}{L(1,\chi)} - \frac{1}{2}\frac{\Gamma'(1)}{\Gamma(1)} + \log\frac{\pi}{k}.$$

Daraus und aus (3) folgt

$$\sum_{\varrho^*}\left(\frac{1}{s-\varrho^*} + \frac{1}{\varrho^*}\right) = \frac{L'(s,\chi)}{L(s,\chi)} + \frac{1}{2}\frac{\Gamma'\left(\frac{s+1}{2}\right)}{\Gamma\left(\frac{s+1}{2}\right)} + \frac{L'(1,\chi)}{L(1,\chi)} + \frac{1}{2}\frac{\Gamma'(1)}{\Gamma(1)} - \log\frac{\pi}{k}. \tag{4}$$

Es sei für das folgende $s = \frac{3}{2} + it$, $\varrho^* = \varrho_1^* + i\varrho_2^*$; t, ϱ_1^* und ϱ_2^* reell. Dann ist wegen $0 < \varrho_1^* < 1$

$$R\left(\frac{1}{\varrho^*}\right) = \frac{\varrho_1^*}{\varrho_1^{*2} + \varrho_2^{*2}} > 0; \qquad R\left(\frac{1}{s-\varrho^*}\right) = \frac{\frac{3}{2} - \varrho_1^*}{\left(\frac{3}{2} - \varrho_1^*\right)^2 + (t - \varrho_2^*)^2} > 0.$$

---

[17]) Landau, Handbuch 1, S. 507.
[18]) Landau, Handbuch 1, S. 511.

$R(\ )$ bedeutet hier den reellen Teil der betreffenden komplexen Größe. Aus den beiden obenstehenden Relationen schließt man weiter

$$(5)\qquad \sum_{|t-\varrho_2^*|\leqq\frac{1}{2}}{}' R\left(\frac{1}{s-\varrho^*}\right) < R\frac{L'(s,\chi)}{L(s,\chi)} + \frac{1}{2}R\frac{\Gamma'\left(\frac{s+1}{2}\right)}{\Gamma\left(\frac{s+1}{2}\right)} + \frac{L'(1,\chi)}{L(1,\chi)} + \frac{1}{2}\frac{\Gamma'(1)}{\Gamma(1)} - \log$$

Im Falle reeller Charaktere liegen die Nullstellen symmetrisch zu $\sigma=\frac{1}{2}$, d. h. mit der Nullstelle $\varrho^*=\varrho_1^*+i\varrho_2^*$ gibt es auch eine $\varrho^{**}=1-\varrho_1^*+i\varrho_2^*$. Es ist

$$R\left(\frac{1}{s-\varrho^*}\right)+R\left(\frac{1}{s-\varrho^{**}}\right)=\frac{\frac{3}{2}-\varrho_1^*}{\left(\frac{3}{2}-\varrho_1^*\right)^2+(t-\varrho_2^*)^2}+\frac{\frac{1}{2}+\varrho_1^*}{\left(\frac{1}{2}+\varrho_1^*\right)^2+(t-\varrho_2)^2}.$$

Diese Funktion von ϱ_1^* hat bei festem t und ϱ_2^* ihren kleinsten Wert im Intervall $(0,1)$ an der Stelle $\varrho_1^*=\frac{1}{2}$, wovon man sich etwa auf folgende Weise überzeugen kann: Es sei $\varrho_1^*=\frac{1}{2}+\alpha$. Dann ist zu zeigen, daß die Funktion von α

$$\Phi(\alpha)=\frac{1-\alpha}{(1-\alpha)^2+(t-\varrho_2^*)^2}+\frac{1+\alpha}{(1+\alpha)^2+(t-\varrho_2^*)^2}$$

ihren kleinsten Wert im Intervall $-\frac{1}{2}\leqq\alpha\leqq+\frac{1}{2}$ annimmt für $\alpha=0$. Dabei soll $|t-\varrho_2^*|\leqq\frac{1}{2}$ sein; es sei abgekürzt $a^2=(t-\varrho_2^*)^2$. Man bilde

$$\Phi(\alpha)-\Phi(0)=\frac{1-\alpha}{(1-\alpha)^2+a^2}+\frac{1+\alpha}{(1+\alpha)^2+a^2}-\frac{2}{1+a^2}=\frac{2\alpha^2\cdot[1-3a^2-\alpha^2]}{[(1-\alpha)^2+a^2]\cdot[(1+\alpha)^2+a^2]\cdot[}$$

Der Nenner ist stets positiv, der Zähler ist $\geqq 2\alpha^2\cdot[\frac{1}{4}-\alpha^2]\geqq 0$ für $|\alpha|\leqq\frac{1}{2}$. Also nimmt $\Phi(\alpha)$ im Intervall $(-\frac{1}{2},+\frac{1}{2})$ tatsächlich seinen kleinsten Wert an für $\alpha=0$. Folglich ist

$$R\left(\frac{1}{s-\varrho^*}\right)+R\left(\frac{1}{s-\varrho^{**}}\right)\geqq 2\cdot\frac{1}{1+a^2}\geqq\frac{2}{1+\frac{1}{4}}=\frac{8}{5}.$$

Nun ist offenbar

$$\frac{1}{2}\cdot\left[N\left(t+\tfrac{1}{2}\right)-N\left(t-\tfrac{1}{2}\right)\right]\cdot\mathrm{Min}\left\{R\left(\frac{1}{s-\varrho^*}\right)+R\left(\frac{1}{s-\varrho^{**}}\right)\right\}\leqq\sum_{|t-\varrho_2^*|\leqq\frac{1}{2}}R\left(\frac{1}{s-\varrho^*}\right.$$

also nach (5), wenn man für das Minimum den gefundenen Wert $\frac{8}{5}$ einsetzt,

$$(6)\quad N\left(t+\tfrac{1}{2}\right)-N\left(t-\tfrac{1}{2}\right)<\frac{5}{4}\cdot\left[R\frac{L'(s,\chi)}{L(s,\chi)}+\frac{1}{2}R\frac{\Gamma'\left(\frac{s+1}{2}\right)}{\Gamma\left(\frac{s+1}{2}\right)}+\frac{L'(1,\chi)}{L(1,\chi)}+\frac{1}{2}\frac{\Gamma'(1)}{\Gamma(1)}-\right.$$

Es sei nun

a) $k = 3$; dann ist

$$R\frac{L'(s,\chi)}{L(s,\chi)} \leqq \left|\frac{L'(s,\chi)}{L(s,\chi)}\right| \leqq \sum_{\nu=1}^{\infty} \frac{a_\nu |\chi(\nu)|}{\nu^{\frac{3}{2}}} < \frac{\log 2}{2^{\frac{3}{2}}} + \frac{\log 2}{4^{\frac{3}{2}}} + \frac{\log 5}{5^{\frac{3}{2}}} + \int_6^{\infty} \frac{\log x}{x^{\frac{3}{2}}}\, dx < 3{,}6;$$

$$L(1,\chi) = 1 - \frac{1}{2} + \frac{1}{4} - \frac{1}{5} + \ldots > \frac{1}{2}, \qquad L'(1,\chi) = \frac{\log 2}{2} - \frac{\log 4}{4} + \ldots < \frac{\log 2}{2},$$

$$\frac{L'(1,\chi)}{L(1,\chi)} < \log 2 < 0{,}7; \qquad -\log\frac{\pi}{k} = -\log\frac{\pi}{3} < 0.$$

Also insgesamt

7) $$R\frac{L'(s,\chi)}{L(s,\chi)} + \frac{L'(1,\chi)}{L(1,\chi)} - \log\frac{\pi}{k} < 4{,}3.$$

b) $k = 4$; dann ist

$$R\frac{L'(s,\chi)}{L(s,\chi)} < \frac{\log 3}{3^{\frac{3}{2}}} + \frac{\log 5}{5^{\frac{3}{2}}} + \int_6^{\infty} \frac{\log x}{x^{\frac{3}{2}}}\, dx < 3{,}55;$$

$$L(1,\chi) = 1 - \frac{1}{3} + \frac{1}{5} \ldots > 0{,}78, \qquad L'(1,\chi) = \frac{\log 3}{3} - \frac{\log 5}{5} + \ldots < \frac{\log 3}{3},$$

$$\frac{L'(1,\chi)}{L(1,\chi)} < \frac{\log 3}{3 \cdot 0{,}78} < 0{,}5; \qquad -\log\frac{\pi}{k} = \log\frac{4}{\pi} < 0{,}25.$$

Also wieder

7) $$R\frac{L'(s,\chi)}{L(s,\chi)} + \frac{L'(1,\chi)}{L(1,\chi)} - \log\frac{\pi}{k} < 4{,}3.$$

Weiter ist

$$\frac{\Gamma'(s)}{\Gamma(s)} = -\frac{1}{s} - C + s \cdot \sum_{n=1}^{\infty} \frac{1}{n \cdot (s+n)} \qquad (C = 0{,}57\ldots).$$

Folglich

8) $$\frac{\Gamma'(1)}{\Gamma(1)} = -1{,}57 + \sum_{n=1}^{\infty} \frac{1}{n \cdot (n+1)} < 0.$$

Und schließlich ist[19])

$$-\frac{\Gamma'(z)}{\Gamma(z)} = \frac{1}{2z} + \int_0^{\infty} \left(\frac{1}{z+u} - \frac{1}{1+u}\right) du + \frac{f_1(0)}{z^2} - 2\int_0^{\infty} \frac{f_1(u)}{(u+z)^3}\, du.$$

Dabei ist $|f_1(u)| < 1$. Setzt man $z = z_1 + i z_2$, dann wird

$$-\frac{\Gamma'(z)}{\Gamma(z)} = \frac{1}{2z} + \frac{f_1(0)}{z^2} - 2\int_0^{\infty} \frac{f_1(u)}{(u+z)^3}\, du - \frac{1}{2}\log(z_1^2 + z_2^2) + i \operatorname{arctg}\frac{z_1}{z_2} - i \cdot \frac{\pi}{2}.$$

[19]) Vgl. Landau, Vorl. üb. Zahlentheorie 2, S. 81.

Hierin ist für $z_1 > 0$

$$2 \cdot \left| \int_0^\infty \frac{f_1(u)}{(u+z)^3}\, du \right| < 2 \cdot \int_0^\infty \frac{du}{(z_2^2+u^2)^{\frac{3}{2}}} = \frac{2}{z_2^2} \cdot \int_0^\infty \frac{dv}{(1+v^2)^{\frac{3}{2}}} < \frac{\pi}{z_2^2}$$

und

$$R\left(\frac{1}{2z}\right) < \frac{z_1}{2z_2^2}.$$

Also ist für $z_2 > 0$

$$R\frac{\Gamma'(z)}{\Gamma(z)} < \frac{1}{2}\log(z_1^2+z_2^2) + \frac{1}{z_2^2}\cdot\left(\pi+1+\frac{z_1}{2}\right) < \log z_2 + \frac{1}{z_2^2}\cdot\left(\frac{z_1^2}{2}+\pi+1+\frac{z_1}{2}\right)$$

und speziell (für $t > 0$)

$$(9) \qquad R\frac{\Gamma'\left(\frac{\frac{3}{2}+it+1}{2}\right)}{\Gamma\left(\frac{\frac{3}{2}+it+1}{2}\right)} = R\frac{\Gamma'\left(\frac{5}{4}+i\frac{t}{2}\right)}{\Gamma\left(\frac{5}{4}+i\frac{t}{2}\right)} < \log\frac{t}{2} + \frac{24}{t^2}.$$

Aus (6), (7), (8) und (9) folgt für positives t

$$N\left(t+\frac{1}{2}\right) - N\left(t-\frac{1}{2}\right) < \frac{5}{4}\cdot\left(\frac{1}{2}\log\frac{t}{2} + \frac{12}{t^2} + 4{,}3\right) < \frac{5}{4}\cdot\left(\frac{1}{2}\log t + 4 + \frac{12}{t^2}\right.$$

oder

$$(10) \qquad N\left(t+\frac{1}{2}\right) - N\left(t-\frac{1}{2}\right) < \frac{5}{8}\log t + 5 + \frac{15}{t^2}.$$

Diese Abschätzung über die Anzahl der Nullstellen gilt also dabei für di Reihen mit den Nicht-Hauptcharakteren nach dem Modul 3 *und* nach den Modul 4.

Außerdem haben beide Reihen noch die trivialen Nullstellen $-1.\ -3\ -5, \ldots$.

χ_1 sei nun der Hauptcharakter mod k (k wieder $= 3$ oder $= 4$) Dann ist[20])

$$L(s, \chi_1) = \left(1 - \frac{1}{p^s}\right)\cdot\zeta(s);$$

dabei ist $p = 3$ für $k = 3$ und $p = 2$ für $k = 4$. $L(s, \chi_1)$ verschwinde daher an den Nullstellen von $\zeta(s)$ und außerdem an den Stellen

$$s = \frac{2m\pi i}{\log p} \qquad (m = 0, \pm 1, \pm 2, \ldots)$$

§ 3.

χ sei wieder der Nicht-Hauptcharakter. Es ist

$$\frac{1}{2\pi i}\int_{(\mathfrak{L})} \frac{x^s}{s^4}\frac{L'(s,\chi)}{L(s,\chi)}\, ds = \sum \text{Resid.},$$

[20]) Landau, Handbuch 1, S. 423.

wobei die Summe erstreckt werden soll über alle Pole des Integranden in der Halbebene $\sigma < 2$. Beweis genau wie in Teil I, § 2, nur mit dem Unterschied, daß jetzt über den Rand des Rechtecks mit den Ecken

$$2 - iT, \quad 2 + iT, \quad -2m + iT, \quad -2m - iT$$

integriert werden muß.

Der Integrand hat Pole an folgenden Stellen:

a) an den Nullstellen ϱ^* von $L(s, \chi)$ im Gebiet $0 < \sigma < 1$. Summe der Residuen ist hier $\sum_{\varrho^*} \frac{x^{\varrho^*}}{\varrho^{*4}}$;

b) an den Stellen $s = -(2n+1)$ $(n = 0, 1, 2, \ldots)$. Summe der Residuen: $\sum_{n=0}^{\infty} \frac{x^{-(2n+1)}}{(2n+1)^4}$;

c) an der Stelle $s = 0$; Residuum: $c_0 + c_1 \log x + c_2 \log^2 x + c_3 \log^3 x$.

Also ist

$$\frac{1}{2\pi i} \int_{(2)} \frac{x^s}{s^4} \frac{L'(s, \chi)}{L(s, \chi)} = \sum_{\varrho^*} \frac{x^{\varrho^*}}{\varrho^{*4}} + \sum_{n=0}^{\infty} \frac{x^{-(2n+1)}}{(2n+1)^4} + c_0 + c_1 \log x + c_2 \log^2 x + c_3 \log^3 x.$$

Weiter ist

$$\frac{L'(s, \chi_1)}{L(s, \chi_1)} = \frac{\zeta'(s)}{\zeta(s)} + \frac{\log p}{p^s - 1}. \tag{12}$$

Wenn ϱ wieder die nicht-trivialen Nullstellen von $\zeta(s)$ durchläuft, dann ist nach Teil I, (5) und (6)

$$\frac{1}{2\pi i} \int_{(2)} \frac{x^s}{s^4} \frac{L'(s, \chi_1)}{L(s, \chi_1)} ds = -x + \sum_{\varrho} \frac{x^{\varrho}}{\varrho^4} + \sum_{n=1}^{\infty} \frac{x^{-2n}}{(2n)^4}$$
$$+ b_0 + b_1 \log x + b_2 \log^2 x + b_3 \log^3 x + \frac{1}{2\pi i} \int_{(2)} \frac{x^s}{s^4} \frac{\log p}{p^s - 1} ds.$$

Dieses zweite Integral ist offenbar ebenfalls gleich der Summe der Residuen eines Integranden in der Halbebene $\sigma < 2$. Denn wenn man wieder integriert über den Rand des Rechtecks mit den Ecken

$$2 - iT, \quad 2 + iT, \quad -S + iT, \quad -S - iT,$$

wo $S > 0$ und T etwa $= \frac{(2m+1)\pi}{\log p}$ sein soll (m positiv ganzzahlig), dann ist für $x > 1$

$$\left| \int_{-S+iT}^{2+iT} \frac{x^s}{s^4} \frac{\log p}{p^s - 1} ds \right| < \frac{x^2 \log p \cdot (S+2)}{T^4},$$

entsprechend $\left|\int_{-S-iT}^{2-iT}\right|$. Beide Teilintegrale gehen also mit wachsendem m gegen 0. Und für den dritten Bestandteil gilt

$$\left|\int_{-S-i\infty}^{-S+i\infty}\frac{x^s}{s^4}\frac{\log p}{p^s-1}\,ds\right| < \frac{\log p}{1-p^{-S}}x^{-S}\left|\int_{-S-i\infty}^{-S+i\infty}\frac{ds}{s^4}\right|;$$

auch dieses Integral geht mit wachsendem S gegen 0.

Der Integrand $\frac{x^s}{s^4}\frac{\log p}{p^s-1}$ hat Pole an den Stellen 0 und $\frac{2\mu\pi i}{\log p}$, $\mu=\pm1,\ \pm2,\ \pm3,\ \ldots$. Die Summe der Residuen an den Stellen $\frac{2\mu\pi i}{\log p}$ ist

$$\sum_{\mu=-\infty}^{+\infty}{}' \frac{x^{\frac{2\mu\pi i}{\log p}}}{\left(\frac{2\mu\pi i}{\log p}\right)^4}.$$

Der Strich besagt dabei, $\mu=0$ soll in der Summierung ausgelassen werden.

Das Residuum an der Stelle $s=0$ endlich erhält man durch Reihenentwicklung des Integranden

$$\frac{x^s}{s^4}\cdot\frac{\log p}{p^s-1}=\frac{\left(1+s\log x+\ldots+\frac{s^4\log^4 x}{4!}+\ldots\right)\cdot\log p}{s^4\cdot(s\log p+\ldots)}.$$

Das Residuum ist also gleich $\frac{\log^4 x}{4!}=\frac{\log^4 x}{24}$. Folglich ist

$$(13)\quad \frac{1}{2\pi i}\int_{(2)}\frac{x^s}{s^4}\frac{L'(s,\chi_1)}{L(s,\chi_1)}\,ds=-x+\sum_{\varrho}{}'\frac{x^\varrho}{\varrho^4}+\sum_{n=1}^{\infty}\frac{x^{-2n}}{(2n)^4}+\sum_{\mu=-\infty}^{+\infty}{}'\frac{x^{\frac{2\mu\pi i}{\log p}}}{\left(\frac{2\mu\pi i}{\log p}\right)^4}$$
$$+b_0+b_1\log x+b_2\log^2 x+b_3\log^3 x+\frac{1}{24}\log^4 x.$$

Aus (2), (11) und (13) folgen dann die expliziten Ausdrücke für $k=3$ oder $k=4$, also $h=\varphi(k)=2$, und für $l=+1$ oder $l=-1$:

$$(14)\quad \frac{1}{3}\sum_{\substack{\nu\leqq x\\ \nu\equiv l\,(\mathrm{mod}\,k)}}a_\nu\log^3\frac{x}{\nu}=x-\sum_{\varrho}{}'\frac{x^\varrho}{\varrho^4}-\sum_{n=1}^{\infty}\frac{x^{-2n}}{(2n)^4}-\sum_{\mu=-\infty}^{+\infty}{}'\frac{x^{\frac{2\mu\pi i}{\log p}}}{\left(\frac{2\mu\pi i}{\log p}\right)^4}\mp\sum_{\varrho^*}\frac{x^{\varrho^*}}{\varrho^{*4}}\mp\sum_{n=0}^{\infty}\frac{x^{-}}{(2}$$
$$-d_0-d_1\log x-d_2\log^2 x-d_3\log^3 x-\frac{1}{24}\log^4 x.$$

Dabei sind die d_i Konstante, deren Wert ohne Bedeutung ist.

Von den rechts stehenden Summen sollen zunächst die zweite, die dritte und die fünfte abgeschätzt werden. Es ist

$$(15)\qquad \left|\sum_{\mu=-\infty}^{+\infty}{}' \frac{x^{\frac{2\mu\pi i}{\log p}}}{\left(\frac{2\mu\pi i}{\log p}\right)^4}\right| \leqq 2\left(\frac{\log p}{2\pi}\right)^4 \sum_{\mu=1}^{\infty}\frac{1}{\mu^4} \leqq 2\left(\frac{\log 3}{2\pi}\right)^4 \sum_{\mu=1}^{\infty}\frac{1}{\mu^4} < 0{,}003;$$

$$(16)\qquad \left|\sum_{n=1}^{\infty}\frac{x^{-2n}}{(2n)^4} \pm \sum_{n=0}^{\infty}\frac{x^{-(2n+1)}}{(2n+1)^4}\right| \leqq \sum_{n=1}^{\infty}\frac{x^{-n}}{n^4} < \frac{1{,}1}{x} \quad \text{für} \quad x > 1.$$

§ 4.

Es sei abgekürzt

$$\beta(x) = \beta(x;\, k, l) = \sum_{\substack{\nu \leqq x \\ \nu \equiv l \,(\mathrm{mod}\, k)}} a_\nu \log^3 \frac{x}{\nu}.$$

Δ soll ganz analog zu Teil I, § 3 gebildet werden:

$$\Delta = \beta((1+y)^4 x) - 4\beta((1+y)^3 x) + 6\beta((1+y)^2 x) - 4\beta((1+y)x) + \beta(x).$$

y bedeute dabei wieder eine reelle positive Größe. Dann ist, entsprechend zu Formel (8) in Teil I,

$$(17)\qquad \Delta \leqq 4y^3 \sum_{\substack{x < \nu \leqq (1+y)^4 x \\ \nu \equiv l \,\mathrm{mod}\, k)}} a_\nu.$$

Setzt man in Δ für die Funktionen β ihren Wert nach (14) ein, dann fallen nach Teil I, (7) die logarithmischen Glieder mit Ausnahme des letzten weg, dieses nimmt den Wert $\frac{1}{24}\cdot 24 \log^4(1+y) < y^4$ an, und man erhält unter Berücksichtigung von (15) und (16):

$$\frac{1}{8} y^3 \sum_{\substack{x < \nu \leqq (1+y)^4 x \\ \nu \equiv l \,\mathrm{mod}\, k)}} a_\nu > y^4 x - \left|\sum_{\varrho} \frac{x^\varrho [(1+y)^\varrho - 1]^4}{\varrho^4}\right| - \left|\sum_{\varrho^*} \frac{x^{\varrho^*}[(1+y)^{\varrho^*}-1]^4}{\varrho^{*4}}\right|$$

$$- y^4 - 16\cdot 0{,}003 - 16\cdot\frac{1{,}1}{x}$$

oder

$$y^3 \sum_{\substack{x < \nu \leqq (1+y)^4 x \\ \nu \equiv l \,(\mathrm{mod}\, k)}} a_\nu > y^4 x - (2+y)^4 \sum_{\varrho}\left|\frac{x^\varrho}{\varrho^4}\right| - (2+y)^4 \sum_{\varrho^*}\left|\frac{x^{\varrho^*}}{\varrho^{*4}}\right| - y^4 - 0{,}05 - \frac{18}{x}.$$

Es sei von jetzt an $y = \frac{1}{5{,}5}$; dann ist $(1+y)^4 < 2$ und $(2+y)^4 < 23$.

§ 5.

Nun gilt für die beiden Nicht-Hauptcharaktere (mod 3 und mod 4), daß die Nullstellen ϱ^* von $L(s, \chi)$ mit $|J(\varrho^*)| < 44$ auf $\sigma = \frac{1}{2}$ liegen;[21])

[21]) Großmann, a. a. O., §§ 7 und 8.

$J(\varrho^*)$ bedeute dabei den Imaginärteil von ϱ^*. Da die Nullstellen symmetrisch liegen zu $\sigma = \frac{1}{2}$, so hat von den übrigen höchstens die Hälfte eine Abszisse $R(\varrho^*) = 1$, für mindestens die Hälfte ist $R(\varrho^*) \leqq \frac{1}{2}$. Und da die Nullstellen auch noch symmetrisch zur reellen Achse liegen, so gilt

$$(19)\quad (2+y)^4\left[\sum_{\varrho}\frac{|x^{\varrho}|}{|\varrho|^4}+\sum_{\varrho^*}\frac{|x^{\varrho^*}|}{|\varrho^*|^4}\right]\leqq 23x\left[\sum_{J(\varrho)\geqq 200}\frac{1}{|\varrho|^4}+\sum_{J(\varrho^*)\geqq 44}\frac{1}{|\varrho^*|^4}\right]$$

$$+23\sqrt{x}\left[\sum_{J(\varrho)\geqq 200}\frac{1}{|\varrho|^4}+\sum_{J(\varrho^*)\geqq 44}\frac{1}{|\varrho^*|^4}+2\sum_{0\leqq J(\varrho)<200}\frac{1}{|\varrho|^4}+2\sum_{0\leqq J(\varrho)<14}\frac{1}{|\varrho^*|^4}\right].$$

Hierin ist nach Teil I (16)

$$\sum_{J(\varrho)\geqq 200}\frac{1}{|\varrho|^4}<3{,}32\cdot 10^{-8}.$$

Weiter ist nach (10)

$$\sum_{J(\varrho^*)\geqq 44}\frac{1}{|\varrho^*|^4}\leqq\sum_{n=44}^{\infty}\frac{N(n+1)-N(n)}{n^4}<\sum_{n=44}^{\infty}\frac{\frac{5}{8}\log\left(n+\frac{1}{2}\right)+5+\frac{15}{n^2}}{n^4}$$

$$<\sum_{n=44}^{\infty}\frac{\frac{5}{8}\log n+5{,}1}{n^4}<\int_{43}^{\infty}\frac{\frac{5}{8}\log x+5{,}1}{x^4}\,dx<3{,}28\cdot 10^{-5}.$$

Also ist

$$(20)\quad 23\left[\sum_{J(\varrho)\geqq 200}\frac{1}{|\varrho|^4}+\sum_{J(\varrho^*)\geqq 44}\frac{1}{|\varrho^*|^4}\right]<23\cdot 3{,}3\cdot 10^{-5}<7{,}6\cdot 10^{-4}.$$

Nach Teil I, (18) ist

$$(21)\quad \sum_{0\leqq J(\varrho)<200}\frac{1}{|\varrho|^4}<\frac{1}{14^4}+\frac{2}{20^4}+\frac{7}{30^4}+\frac{19}{50^4}+\frac{50}{100^4}<5{,}1\cdot 10^{-5}.$$

Die Nullstellen von $L(s,\chi)$ mit $0\leqq J(\varrho)<44$ liegen in den beiden Fällen folgendermaßen:[22])

a) $k=3$.

$$8\leqq J(\varrho^{*(1)})\leqq 10\leqq J(\varrho^{*(2)})\leqq 12\leqq J(\varrho^{*(3)})\leqq 18\leqq J(\varrho^{*(4)})\leqq 20$$
$$\leqq J(\varrho^{*(5)})\leqq 24\leqq J(\varrho^{*(6)}) \text{ bis } J(\varrho^{*(8)})\leqq 30\leqq J(\varrho^{*(9)}) \text{ bis } J(\varrho^{*(14)})\leqq 44.$$

b) $k=4$.

$$6\leqq J(\varrho^{*(1)})\leqq 10\leqq J(\varrho^{*(2)})\leqq 12\leqq J(\varrho^{*(3)})\leqq 16\leqq J(\varrho^{*(4)})\leqq 18$$
$$\leqq J(\varrho^{*(5)})\leqq 20\leqq J(\varrho^{*(6)}) \text{ und } J(\varrho^{*(7)})\leqq 24$$
$$\leqq J(\varrho^{*(8)}) \text{ bis } J(\varrho^{*(10)})\leqq 30\leqq J(\varrho^{*(11)}) \text{ bis } J(\varrho^{*(17)})\leqq 45.$$

---

[22]) Großmann, a. a. O., §§ 7 und 8.

Also ist für $k=3$

$$\sum_{0\leqq J(\varrho^*)<44}\frac{1}{|\varrho^*|^4}\leqq\frac{1}{8^4}+\frac{1}{10^4}+\frac{1}{12^4}+\frac{1}{18^4}+\frac{1}{20^4}+\frac{3}{24^4}+\frac{6}{30^4}<4{,}3\cdot10^{-4}$$

und für $k=4$

$$\sum_{0\leqq J(\varrho^*)<44}\frac{1}{|\varrho^*|^4}\leqq\frac{1}{6^4}+\frac{1}{10^4}+\frac{1}{12^4}+\frac{1}{16^4}+\frac{1}{18^4}+\frac{2}{20^4}+\frac{3}{24^4}+\frac{7}{30^4}<9{,}76\cdot10^{-4}.$$

Hieraus und aus (21) folgt für beide Fälle

$$(22)\quad 2\cdot23\left[\sum_{0\leqq J(\varrho)<200}\frac{1}{|\varrho|^4}+\sum_{0\leqq J(\varrho^*)<44}\frac{1}{|\varrho^*|^4}\right]$$
$$<46\,[5{,}1\cdot10^{-5}+97{,}6\cdot10^{-5}]<0{,}048.$$

Also ist nach (18), (19), (20) und (22)

$$\frac{1}{5{,}5^3}\sum_{\substack{x<\nu\leqq2x\\ \nu\equiv l\,(\mathrm{mod}\,k)}}a_\nu>x\left[\frac{1}{5{,}5^4}-7{,}6\cdot10^{-4}\right]-\sqrt{x}\,[7{,}6\cdot10^{-4}+0{,}048]-0{,}05-\frac{1}{5{,}5^4}-\frac{18}{x}$$
$$>3{,}3\cdot10^{-4}x-0{,}049\sqrt{x}-0{,}06-\frac{18}{x}.$$

Es sei nun $x\geqq10^6$; dann ist

$$\frac{4}{3}\cdot\frac{1}{5{,}5^3}\sum_{\substack{x<\nu\leqq2x\\ \nu\equiv l\,(\mathrm{mod}\,k)}}a_\nu>(3{,}3-0{,}5)\cdot10^{-4}x=2{,}8\cdot10^{-4}x$$

oder

$$(23)\qquad\sum_{\substack{x<\nu\leqq2x\\ \nu\equiv l\,(\mathrm{mod}\,k)}}a_\nu>\frac{2{,}8\cdot5{,}5^3\cdot3}{4}\cdot10^{-4}x>3{,}4\cdot10^{-2}x.$$

Weiter ist nach Definition

$$\sum_{\substack{x<\nu\leqq2x\\ \nu\equiv l\,(\mathrm{mod}\,k)}}a_\nu-\sum_{\substack{x<p\leqq2x\\ p\equiv l\,(\mathrm{mod}\,k)}}\log p=\sum_{\substack{x<p^\mu\leqq2x\\ p^\mu\equiv l\,(\mathrm{mod}\,k)\\ \mu>1}}\log p\leqq\sum_{\substack{p^\mu\leqq2x\\ \mu>1}}\log p$$
$$=\psi(2x)-\vartheta(2x)\leqq2\psi(\sqrt{2x}),$$

also für $x\geqq10^6$

$$\sum_{\substack{\nu\leqq2x\\ (\mathrm{mod}\,k)}}a_\nu-\sum_{\substack{x<p\leqq2x\\ p\equiv l\,(\mathrm{mod}\,k)}}\log p<2\,[1{,}106\sqrt{2x}+3\log^2\sqrt{2x}+8\log\sqrt{2x}+5]<3{,}8\sqrt{x}.$$

Daraus endlich und aus (23) folgt

$$(24)\qquad\sum_{\substack{x<p\leqq2x\\ p\equiv l\,(\mathrm{mod}\,k)}}\log p>3{,}4\cdot10^{-2}x-3{,}8\sqrt{x}>3\cdot10^{-2}x>0,$$

für $x\geqq10^6$, $k=3$ und $l=1$ oder $k=3$ und $l=2$ oder $k=4$ und $l=1$ oder $k=4$ und $l=3$. D. h. für $x\geqq10^6$ liegen zwischen x und $2x$ stets Primzahlen einer jeden der vier Progressionen.

Um den Satz für $7 \leqq x < 10^6$ zu beweisen, braucht man wieder Primzahltabellen. Für $24 \leqq x < 10^6$ liegen zwischen x und $2x$ sogar stets Primzahlen einer jeden der beiden Progressionen $12n+1$ und $12n-1$; denn in den im folgenden angeführten zwei Systemen von Primzahlen der Form $12n+1$ bzw. $12n-1$ ist die jeweils erste kleiner als $2 \cdot 24$, jede der folgenden kleiner als das Doppelte der vorhergehenden und die letzte größer als 10^6:

a) $12n+1$.

37, 73, 109, 193, 373, 733, 1453, 2857, 5701, 11353, 22669, 45289, 90529, 180949, 361873, 723661, 1447309.

b) $12n-1$.

47, 83, 131, 251, 491, 971, 1823, 3623, 7211, 14411, 28751, 57467, 114827, 229631, 459167, 918263, 1836479.

Und für $7 \leqq x < 24$ endlich liegen zwischen x und $2x$ stets Primzahlen einer jeden der vier Progressionen, wie die folgenden vier Systeme von Primzahlen zeigen:

a) $3n+1$: 13, 19, 37
b) $3n+2$: 11, 17, 29
c) $4n+1$: 13, 17, 29
d) $4n+3$: 11, 19, 31.

(Eingegangen am 15. September 1930.)

Neue Grundlagen der Logik und Mathematik.

Zweite Mitteilung.

Von

Leon Chwistek in Lwów.

Die eingehende Revision der Grundbegriffe der Semantik und der Logik, die ich auf Anregung von Herrn Hetper unternommen habe und an welcher die Herren Hetper und Herzberg erfolgreich mitgearbeitet haben, hat eine unerwartete Vereinfachung der von mir gegebenen Grundlagen der apriorischen Wissenschaften[1]) herbeigeführt. Der Umschwung der Fundamentalideen, der dadurch verursacht wurde, scheint dazu geeignet zu sein, die reine (verzweigte) Typenlehre auf eine einfache typenartige Semantik zurückzuführen[2]), wodurch das System der apriorischen Wissenschaften eine monumentale Einheit erreichen dürfte. Über diese Arbeiten möchte ich hier kurz berichten.

I. Elementare Semantik.

1. Eine wesentliche Schwierigkeit des Systems der elementaren Semantik stammt davon, daß c kein eigentlicher Ausdruck ist. Nun habe ich gezeigt, daß es eine derartige Interpretation des Schemas $*EFGHJ$ gibt, daß c als ein eigentlicher Ausdruck behandelt werden kann. Diese Bemerkung ist recht fruchtbar. Sie erlaubt uns zunächst, den Sinn des Fundamentalschemas der Logik und Mathematik insofern zu verallgemeinern, daß wir es mit einem einzigen Fundamentalbegriff der Logik und Mathematik zu tun haben werden. Ich sage einfach, daß in allen Fällen, wo $*EFGHJ$ als ein logischer Ausdruck auftritt, die Interpretation „für alle E gilt ent-

[1]) Neue Grundlagen der Logik und Mathematik, Math. Zeitschr. **30** (1929), S. 704—724.

[2]) Die Grundlagen der theoretischen Semantik sind von mir in dem Aufsatz „Une méthode métamathématique d'analyse" (Comptes rendus du 1er congrès des mathématiciens des pays slaves, Warszawa 1929) kurz angegeben worden.

weder $\sim F_G\{H\}$ oder $\sim J$" benützt werden kann, was zu folgender Äquivalenz zwischen der neuen und alten Schreibweise führt:

$$*EFGHJ \equiv (E).F_G\{H\} \mid J.$$

Nun wollen wir uns nur auf die Fälle beschränken, in denen entweder E gleich c ist oder E, G, H und F, J voneinander nicht verschieden sind. Wir haben es also nur mit den Formen $*cFGHJ$ und $*EFEEF$ zu tun. Fassen wir die Regel der Transformation dermaßen zusammen, daß $(E)..F_E\{E\} \mid J$ äquivalent $(E).F \mid J$ sei, und setzen wir noch

$$.(c)..F_G\{H\} \mid J \equiv ..F_G\{H\} \mid J,$$

so ist es klar, daß wir die üblichen Begriffe der Logik aus unserem Begriffe bekommen können.

Ich möchte noch folgende Äquivalenzen hervorheben:

$$..F_M\{E\} \equiv \sim *cFME \supset pp,$$

$$.p \equiv \sim *cpcc \supset pp.$$

Aus der zweiten Äquivalenz ist zu ersehen, daß jetzt der Begriff $.p_c\{c\}$ gilt, und zwar einfach als p zu interpretieren ist.

Hat man diese Begriffe einmal gewonnen, so liegt kein Grund vor, irgendeinen formalen Unterschied zwischen $(x).f_a\{x\}$ und $\hat{x}[.f_a\{x\}]$ bzw. zwischen $\prod x.f_a\{x\}$ und $\prod x.f_a\{x\}[a]$ zu machen. Wir können einfach sagen, daß die Menge $\hat{x}[.f_a\{x\}]$ durch den Quantifikatorausdruck $\prod x.f_a\{x\}$ gegeben ist. Nun hat Dr. Greniewski bemerkt, daß dadurch der „Wahrheitssatz":

$$..p \equiv q \supset ..F\{p\} \equiv .F\{q\}$$

falsch wird, da ja aus $.\prod x.f_a\{x\} \equiv \prod x.g_a\{x\}$ das Urteil $..f_a\{y\} \equiv .g_a\{y\}$ keineswegs folgt. Diese Bemerkung hat mich seinerzeit bewogen, von dem hier angegebenen Ideengang abzusehen, da ich darin eine Beschränkung des Systems sah. Das ist aber keineswegs der Fall, da es sich hier einfach um die Hinzufügung einer Funktion handelt, die keine Wahrheitsfunktion ist.

Nun ist die Wittgensteinsche Wahrheitstheorie zwar heutzutage Mode, sie ist aber so oder so mit der Grundidee der Semantik im Widerspruch. Es liegt also für uns kein Grund vor, mit ihr zu rechnen.

2. Einer geistreichen Bemerkung von Herrn Hetper nach können wir von dem Begriff eines elementaren Urteilsausdrucks gänzlich absehen. In der elementaren Semantik, wo es ja keine Mengen von Ausdrücken gibt, kann ein beliebiger eigentlicher Ausdruck als ein elementarer Urteilsausdruck betrachtet werden. Was die Quantifikatorausdrücke anbelangt, so müssen sie offenbar auf Grund einer speziellen Regel eingeführt werden. Wir haben nun folgende Regeln über Buchstaben und eigentliche Ausdrücke

0·01. Buchstaben.

0·011. $c\,\xi\,\eta\,\zeta\,\theta\,\iota\,\lambda\,\mu\,\nu\,\varrho\,\omega\,d\,h\,l\,m\,n$ sind *Buchstaben.*

0·012. $\xi\,\eta\,\zeta\,\theta\,\iota\,\lambda\,\mu\,\nu\,\varrho\,\omega\,d\,h\,l\,m\,n$ sind *semantische Buchstaben.*

0·013. $\mathfrak{a}\,\mathfrak{c}\,\mathfrak{d}\,\mathfrak{e}\,\mathfrak{g}\,\mathfrak{h}\,\mathfrak{i}\,\mathfrak{k}\,\mathfrak{l}\,\mathfrak{m}\,\mathfrak{n}\,\mathfrak{p}\,\mathfrak{q}\,\mathfrak{r}\,\mathfrak{s}\,\mathfrak{t}\,\mathfrak{u}\,\mathfrak{v}\,\mathfrak{w}\,\mathfrak{x}\,\mathfrak{y}\,\mathfrak{z}$ sind *scheinbare Veränderliche.*

0·02. Eigentliche Ausdrücke.

0·021. Ist E ein Buchstabe, so ist E ein *eigentlicher Ausdruck.*

0·022. Sind $E\,F\,G\,H\,J$ eigentliche Ausdrücke und ist c in E enthalten, so ist $*\,E\,F\,G\,H\,J$ ein eigentlicher Ausdruck.

0·023. Sind $E\,F\,G\,H\,J$ eigentliche Ausdrücke und ist E in J nicht enthalten, so ist $*\,E\,F\,G\,H\,J$ ein eigentlicher Ausdruck.

0·024. Ist G ein eigentlicher Ausdruck, J ein semantischer Buchstabe, K eine scheinbare Veränderliche, ist J in G enthalten und ist K in G nicht enthalten und ist F das Resultat der Substitution von K für J in G, so ist $*\,E\,F\,E\,E\,F$ ein eigentlicher Ausdruck.

Wir sehen, daß die elementaren Buchstaben hier gänzlich ausgeschaltet worden sind. Sie werden durch semantische Buchstaben ersetzt werden. Die Axiome 0·5 fallen offenbar weg, die elementaren Verfahrungsregeln werden, wie wir unten sehen werden, wesentlich vereinfacht.

3. Die Abkürzungen sind in der ersten Mitteilung irrtümlicherweise bloß für eigentliche Ausdrücke gegeben worden. Nun handelt es sich offenbar um eine viel allgemeinere Art von Ausdrücken, welche als *semantische Ausdrücke* gelten gelassen werden und welche durch folgende Regeln gegeben sind:

0·03. Abkürzungen.

0·03001. Ist E ein eigentlicher Ausdruck oder eine scheinbare Veränderliche, so ist E ein *semantischer Ausdruck.*

0·03002. Sind $E\,F\,G\,H\,J$ semantische Ausdrücke, so ist $*\,E\,F\,G\,H\,J$ ein semantischer Ausdruck:
Sind $E\,F\,G\,H\,J\,K\,L\,M\,N$ semantische Ausdrücke, so gilt die Tafel der Abkürzungen.

4. Ich gehe nun dazu über, die peinliche Beschränkung $\sim\{.E\,E\}$, wo E ein eigentlicher Buchstabe ist, wegzuschaffen. Bemerken wir, daß in dem Fundamentalschema der Logik und Mathematik der Buchstabe c auf der letzten Stelle nie auftritt, so können wir die Ausdrücke von der Form $*\,E\,F\,G\,H\,c$ auf beliebige Weise interpretieren.

Wir wollen nun zunächst die Abkürzung $.E$ von $*\,E\,c\,c\,c\,c$ einführen, wobei wir von der in der ersten Mitteilung gegebenen Abkürzung $,E$ gänzlich absehen. Nun wird $.E$ als ein Hilfsausdruck betrachtet werden, welcher zur Konstruktion des semantischen Fundamentalschemas und der semantischen Begriffe, insbesondere aber der logischen Buchstaben benützt werden,

direkt aber weder als logischer Buchstabe, noch als ein anderer Begriff der Semantik vorkommen soll.

Als das Fundamentalschema der Semantik lassen wir nun $*EFGH.J$ gelten.

Die Tafel der Abkürzungen wird nun folgendermaßen modifiziert werden:

Wir haben zunächst

Nr.	Abkürzung	Ausdruck	Nr.	Abkürzung	Ausdruck	Nr.	Abkürzung	Ausdruck
0·0303	a	$*\ .ccc.c.c$	0·03031	$,a$	$*\ .ccc..c.c$	0·03032	$„a$	$*\ .ccc...c.$
0·0304	b	$*\ ..ccc.c.c$	0·03041	$,b$	$*\ ..ccc..c.c$	0·03042	$„b$	$*\ ..ccc...c.$
0·0305	e	$*\ ...ccc.c.c$	0·03051	$,e$	$*\ ...ccc..c.c$	0·03052	$„e$	$*\ ...ccc...c.$
0·0306	f	$*....ccc.c.c$	0·03061	$,f$	$*....ccc..c.c$	0·03062	$„f$	$*....ccc...c.$

Die Art und Weise der Konstruktion ist klar. Man kann sie ohne weiteres für die sonstigen eigentlichen Buchstaben der ersten Mitteilung benützen. Es ist zu bemerken, daß die allgemeine Abkürzung $,E$ hier ohne Sinn ist. Die bestrichenen Buchstaben haben n Striche, wenn $n+1$ Punkte auf der vierten Stelle des Ausdrucks auftreten. Der Unterschied zwischen zwei Buchstaben, welche keinen Strich oder dieselbe Anzahl von Strichen aufweisen, wird durch die Anzahl von Punkten auf der ersten Stelle markiert. Die Ausdrücke $.c$, $..c$, $...c$ usw. werden als Typusbuchstaben gelten.

Ich gehe nun zu der Konstruktion weiterer semantischer Begriffe über. Wir haben:

Nr.	Abkürzung	Ausdruck
0·034	$\vert= E$	$*Ec.cc.c$
0·0341	$\vert- E$	$*E.ccc.c$
0·0352	$(EFGH)$	$*EFGH..c$
0·0353	$\breve{\in}EF$	$(EccF)$
0·0357	Expr E	$\vee *Eccc.c\vee\breve{\in}E*cfaxp\breve{\in}E*cpccq$

Der Begriff der semantischen Ordnung $\prec EF$ erscheint jetzt als vollkommen entbehrlich.

Um nun den Begriff Lit E zu gewinnen, wollen wir zunächst den Begriff eines Typusbuchstaben semantisch auffassen. Wir schreiben:

Nr.	Abkürzung	Ausdruck	Sprachliche Übersetzung
0·0355	Lit typ E	$\wedge\Pi u\supset\{Eu\}\vee\breve{\in}uc\exists\mathfrak{y}\breve{\in}u.\mathfrak{y}\sim\breve{\in}Ec$	E ist ein Typusbuchstabe

Dann haben wir die folgende Interpretation für Lit E:

Nr.	Abkürzung	Ausdruck
0·0356	Lit E	$\exists\mathfrak{w}\wedge$ Lit typ $\mathfrak{w}\exists\mathfrak{v}\wedge\breve{\in}E*\mathfrak{v}cc\mathfrak{w}.c$ Lit typ $\mathfrak{v}$

Weiter sind in der Tafel der Abkürzungen folgende Modifikationen durchzuführen:

Nr.	Abkürzung	Ausdruck
0·0358	Funct E	$\wedge$ Expr $E \exists a \exists c \overline{=}_{\mathfrak{o}} E * a c a a c$
0·0359	Prop E	Expr E
0·0369	Cl $E J F M$	$\wedge$ Expr $E \wedge$ arg $E J M \exists a \wedge \sim \{E a\} \wedge$ Lit $a \exists b$ $\wedge \overline{=}_{\mathfrak{o}} F \Pi a b (E J a b)$
0·0384	D (gen) $G H$	$\exists i \exists \mathfrak{k} \exists [\exists m \exists n \exists \xi \wedge$ Cl $[i [\xi \wedge \sim \{m i\}$ $\wedge$ Expr $\mid \mathfrak{k} \mid m \mid [n \wedge \overline{=}_{\mathfrak{o}} G \wedge \supset m [\supset [n \overline{=}_{\mathfrak{o}} H \wedge \supset m \mathfrak{k} \supset \mathfrak{k} n$

Hier ist die letzte Regel mehr intuitiv als es in der ersten Mitteilung der Fall ist.

Der Begriff Gen $E J F$ ist jetzt einfach zu streichen.

Ich möchte noch bemerken, daß in der ersten Mitteilung der Begriff compat $E F$ irrtümlicherweise zu enge gefaßt wurde. Statt der Bedingung $\wedge$ typ $E m \wedge$ typ $F m \{m \mathfrak{l}\}$ ist zu lesen: $\wedge \vee$ typ $E m \overline{=}_{\mathfrak{o}} E m \wedge \vee$ typ $F m$ $\overline{=}_{\mathfrak{o}} F m \{m \mathfrak{l}\}$.

5. Die elementaren Verfahrungsregeln sind jetzt die folgenden:

Modus ponens.

0·11. Haben wir $\models E$ und $\models \mid E \mid G H$, so haben wir $\models H$.

Substitution.

0·121. Haben wir $\models G$ und ist E ein semantischer Buchstabe, sind F, H eigentliche Ausdrücke und ist H das Resultat der Substitution von F für E in G, so haben wir $\models H$.

Verallgemeinerung.

0·13. Haben wir $\models \wedge \supset M G \supset G N$ und ist H das Resultat der Substitution von F für E in G, ist F eine scheinbare Veränderliche und E ein semantischer Buchstabe, welcher in G vorkommt, kommt E in M nicht vor und kommt F in G nicht vor, so haben wir $\models \wedge \supset M \Pi F H \supset \Pi F H N$.

6. Es ist nun klar, daß die Tafel der Axiome auf eine radikale Weise vereinfacht werden kann. Die Ordnungsaxiome fallen ebensowohl wie die Axiome der semantischen Urteilsausdrücke weg. Wir haben zunächst:

0·6. Logische Ausdrücke:

0·61. $\models \supset \wedge$ Expr $\lambda D \lambda \mu$ Expr μ

0·62. $\models \supset$ Cl $\lambda \mu \nu \varrho$ Expr ν

Die sonstigen Axiome beziehen sich auf den Begriff $(EFGH)$. Was nun die Axiomatik dieses Begriffs anbelangt, so kann sie von zwei verschiedenen Standpunkten aus studiert werden. Zunächst kann man die vollständige Axiomatik dieses Begriffes ins Auge fassen. Wir haben dann vor uns ein recht schwieriges Problem, das bisher noch nicht gelöst ist. Nun kann man aber auch daran denken, die Axiome der ersten Mitteilung auf ein Minimum zu reduzieren, d. h. nur so viel annehmen, was zur Konstruktion des Systems notwendig ist. Dieser Ideengang führt zu folgender Axiomentafel, deren letzte Redaktion von Herrn Herzberg verfertigt wurde. Dabei wollen wir die Theorie von Herrn Hetper in dem Sinne benützen, daß wir zwei beliebige identische Ausdrücke als äquivalente Urteile betrachten werden. Wir haben nun

0·7. Semantische Identität.

0·701. $\models \supset =_{\circ} \lambda\mu \equiv \lambda\mu$

0·702. $\models \sim =_{\circ} c * \lambda\mu\nu\xi\eta$

0·73. $\models \supset \wedge(\lambda\xi\eta\mu)(\lambda\xi\eta\nu) =_{\circ} \mu\nu$

0·8. Substitution.

0·81. $\models(\lambda\xi\xi\lambda)$

0·82. $\models(\lambda\lambda\eta\eta)$

0·83. $\models \supset \sim =_{\circ} \xi * \lambda\mu\nu\zeta\theta \equiv (* \lambda\mu\nu\zeta\theta\xi\eta * lmnhd) \wedge (\lambda\xi\eta l) \wedge (\mu\xi\eta m) \wedge (\nu\xi\eta n) \wedge (\zeta\xi\eta h)(\theta\xi\eta d)$.

Hier ist zu bemerken, daß in der ersten Mitteilung die Hypothesen: $\sim =_{\circ} \xi * c\lambda cc\mu$, $\sim =_{\circ} \xi * c\lambda\theta\zeta\mu$, $\sim =_{\circ} \xi * \zeta\lambda\theta\zeta\lambda$ in 0·83 bzw. 0·84 bzw. 0·85 irrtümlicherweise weggelassen wurden.

0·9. Subsumption.

0·901. $\models \vee =_{\circ} c\lambda(c\lambda cc)$

0·92. $\models \vee \{\lambda\nu\}(\lambda\nu\xi\lambda)$

Das Axiom von Nicod muß jetzt mit semantischen Buchstaben geschrieben werden. Wir haben also:

0·3. $\models \| \lambda | \mu\nu | \supset \xi\xi \supset | \eta\mu | \lambda\eta$.

In der ersten Mitteilung ist statt $\supset | \mathfrak{r}\mathfrak{q} | \mathfrak{p}\mathfrak{r}$ offenbar $\supset | \mathfrak{s}\mathfrak{q} | \mathfrak{p}\mathfrak{s}$ zu lesen.

II. Semantische Typenlehre.

7. Fragen wir nun, ob es etwa nicht möglich wäre, die Theorie von Hetper, welche uns einen beliebigen Ausdruck als einen elementaren Urteilsausdruck betrachten läßt, dermaßen zu verallgemeinern, daß sie auch auf die Logik und Mathematik angewandt werden könnte. — Rechnet man damit, daß die bloße Substitution für ein Funktionszeichen ganz passiv ist

und erst durch die Anwendung der Regel der Transformation fruchtbar wird, so ist kein Grund da, irgendeine Beschränkung der rein formalen Substitution einzuführen, welche keine bloße Berücksichtigung des Einklangs der Ausdrücke wäre. Sind die benützten Ausdrücke im Einklang, so bleibt es noch die Fälle zu unterscheiden, in welchen die Substitution *wesentlich* oder etwa *tot* ist. Im ersten Falle werden wir die Regel der Transformation ohne weiteres anwenden können, im anderen Falle muß sie aber aufs neue formuliert werden. Betrachten wir z. B. das Urteil

$$.\hat{x}[.f_a\{\hat{x}\}]\{z\}$$

und setzen für z irgendeinen nicht zulässigen Ausdruck ein, z. B. $\hat{x}[.f_a\{\hat{x}\}]$. Wir werden sagen, daß die Substitution tot ist, und wir werden unsere Transformationsregel derart formulieren, daß

$$..\hat{x}[.f_a\{\hat{x}\}]_a\{\hat{x}[.f_a\{\hat{x}\}]\} \equiv \hat{x}[.f_a\{\hat{x}\}].$$

Im allgemeinen wird für jede tote Substitution von K für J in $.F_E\{J\}$ bzw. in $.J_E\{F\}$ die ausgeartete Transformationsregel:

$$..F_E\{K\} \equiv F \quad \text{bzw.}$$
$$..K_E\{F\} \equiv K$$

gelten. Der zweite Fall ist z. B. durch

$$.\supset p.p_a\{z\} \equiv \supset pp$$

zu illustrieren. —

8. Diese Theorie erlaubt uns den Mengenbegriff auf den Begriff eines Ausdrucks zu reduzieren. Wir haben bloß mit Mengen von Ausdrücken zu tun. Einerseits sind unsere Mengen gewisse Ausdrücke. Andererseits kann ein beliebiger Ausdruck, also auch eine beliebige Menge als ein Urteil interpretiert werden. Wahre Urteile sind volle Mengen. Nicht volle Mengen werden durch falsche Urteile eingeführt. Ein falsches Urteil, das kein Mengenausdruck ist, ist stets eine leere Menge.

Nun kann eine Menge von Ausdrücken keineswegs zu ihrem echten Bereiche gehören. Das macht eine Hierarchie von Ausdrücken unentbehrlich. Wir werden sagen, daß ein Ausdruck M dem semantischen Typus E angehört, wenn E ein Typusbuchstabe ist, also etwa $.c$, $..c$, $...c$ usw., und wenn E in A entweder gar nicht vorkommt oder etwa als Bestandteil eines eigentlichen Buchstabens. Der in einem eigentlichen Buchstaben E auf vierter Stelle vorkommende Typusbuchstabe wird als der Grad von E gelten. Als Typuszeichen werden Typusbuchstaben benützt werden. Die Ausdrücke, welche dem semantischen Typus eines Typuszeichens angehören, bilden seinen echten Bereich. Für den echten Bereich ist die übliche, sonst die ausgeartete Transformationsregel zu benützen. Es ist ratsam,

als Argumente, welche dem Typuszeichen K angehören, eigentliche Buchstaben vom Grade K bzw. von einem höheren Grade zu benützen. Dieselbe Regel gilt bei der Einführung der scheinbaren Veränderlichen. Obwohl nun unseren Mengen beliebige Ausdrücke als Argumente zukommen können, so ist es doch ratsam, in der Logik und Mathematik die Regel der Substitution auf eigentliche Buchstaben und logische Ausdrücke zu beschränken.

9. Die hier kurz charakterisierte semantische Typenlehre steht der Theorie sehr nahe, welche Poincaré zur Klärung der Richardschen Antinomie skizziert hatte. Sie macht aus der Logik und Mathematik ein Kapitel der rationalen Semantik. Will man nun zum Studium dieser Disziplin in ihrem Ganzen übergehen, so wird man den Weg einschlagen müssen, den ich in der oben zitierten Arbeit: Une méthode métamathématique d'analyse angedeutet habe. Jedenfalls wird man durch die Benützung der semantischen Typenlehre eine wesentliche Vereinfachung des Problems erzielen.

(Eingegangen am 19. November 1930.)

Die erste Randwertaufgabe der allgemeinen selbstadjungierten elliptischen Differentialgleichung zweiter Ordnung im Raum für beliebige Gebiete*).

Von

Werner Püschel in Hannover.

Einleitung.

Bekanntlich ist die erste Randwertaufgabe der Potentialtheorie im Raume nicht immer in dem strengen Sinne lösbar, der von der Lösung die Stetigkeit bis in den Rand hinein fordert[1]). Mildert man diese Bedingung jedoch, so läßt sich jedem beliebigen Gebiet und jeder stetigen Randwertvorgabe eine Lösung im erweiterten Sinne zuordnen[2]). Sie ist eindeutig bestimmt durch die Forderung, Grenzfunktion von Folgen gewisser benachbarter Lösungen zu sein. In mehreren Arbeiten wurde die wichtige Frage behandelt, inwieweit die Lösung im erweiterten Sinne die vorgegebenen Randwerte annimmt[3]).

Im folgenden werden die entsprechenden Verhältnisse bei der allgemeinen selbstadjungierten elliptischen Differentialgleichung

$$\Delta_3 u + a(x_1, x_2, x_3)\, u + b(x_1, x_2, x_3) = 0 \tag{1}$$

untersucht; hierbei bezeichnet $\Delta_3 u$ einen elliptischen Differentialausdruck der Form

$$\sum_{i,k=1}^{3} \frac{\partial}{\partial x_i}\left(a_{i,k} \frac{\partial u}{\partial x_k}\right) \qquad (a_{i,k} = a_{k,i}).$$

*) Von der mathematisch-naturwissenschaftlichen Fakultät der Universität Göttingen als Inaugural-Dissertation zugelassen.

[1]) H. Lebesgue, C. R. séances Soc. math. France **41** (1912), S. 17.

[2]) N. Wiener, a) Certain Notions in Potential Theory, Journ. math. phys. Mass. Institut **3** (1924), S. 24—51; b) The Dirichlet Problem, ebenda, S. 127—146.

[3]) Aus ihrer großen Zahl seien folgende hervorgehoben: N. Wiener, vorige Anmerkung. O. D. Kellog, a) On the Classical Dirichlet Problem for General Domains, Proc. Nat. Acad. **12** (1926), S. 397—406; b) Foundations of Potential Theory, Berlin 1929, S. 315—338. Daselbst auch Angaben über weitere Literatur. Ferner F. Vasilesco, Sur les singularités des fonctions harmoniques, Journ. math. pures et appl **9** (9) (1930), S. 81—111.

Das Hauptergebnis ist folgendes: Die erste Randwertaufgabe für beliebige beschränkte Gebiete und stetige Randwerte besitzt stets eine Lösung im erweiterten Sinne[4]). Dabei zerfallen die Randpunkte unabhängig von der speziellen Wahl der Differentialgleichung (1) und der Randwertvorgabe in drei Klassen: die eine Klasse A wird von den *regulären Punkten* gebildet, d. h. von solchen Randpunkten, in denen bei beliebiger Annäherung von innen her die Lösung stets gegen den vorgegebenen Randwert konvergiert. Die zweite Klasse B besteht aus denjenigen nichtregulären Punkten, in denen bei Annäherung auf einer passenden Teilfolge von innen her die Lösung gegen den vorgegebenen Randwert konvergiert. Dagegen haben die auf den Punkten der dritten Klassen C vorgegebenen Randwerte keinerlei Einfluß auf die Lösung, die in diesen Punkten stetig definiert werden kann und dann sogar regulär, d. h. zweimal stetig differenzierbar in ihnen ist. Die Zugehörigkeit eines Randpunktes p zu einer der drei Klassen wird durch das Konvergenzverhalten gewisser von N. Wiener eingeführter Reihen bestimmt, die den einzelnen Randpunkten zugeordnet werden; p gehört zu A, B oder C, je nachdem die zugehörige Reihe divergiert, konvergiert ohne abzubrechen oder aber abbricht.

Der Gang der Untersuchung ist in kurzem folgender: Zunächst wird die Lösung im erweiterten Sinne konstruiert, wonach in den Paragraphen 2 und 3 einige Hilfsmittel für die Untersuchung des Randes bereitgestellt werden. Die Klasse C läßt sich dann sehr einfach erledigen (§ 4). Hierbei ergibt sich insbesondere eine Charakterisierung derjenigen Mengen, welche nur hebbare Singularitäten tragen können. Im nächsten Paragraphen wird der übrige Rand behandelt. Während wir uns in den Paragraphen 1 bis 5 auf Differentialgleichungen der Form $\Delta_2 u = 0$ beschränken, erfolgt die Übertragung der Ergebnisse auf die allgemeine Differentialgleichung im letzten Paragraphen. Dies gelingt ohne Schwierigkeit unter Benutzung einer Integralgleichung. An vielen Stellen lassen sich die Überlegungen aus der Potentialtheorie fast wörtlich übertragen; dort werden wir uns dementsprechend kurz fassen.

Für die Anregung zu dieser Arbeit habe ich Herrn R. Courant zu danken.

§ 1.

Existenz und Eindeutigkeit der Lösung im erweiterten Sinne von $\Delta_2 u = 0$ für beliebige Gebiete.

1. Vorbemerkungen.

Von den Koeffizienten des Differentialausdruckes $\Delta_2 u$ setzen wir zweimalige stetige Differenzierbarkeit voraus, wobei die zweiten Ableitungen

[4]) Falls a und b nicht auftreten, lassen wir auch unbeschränkte Gebiete zu, deren Rand im Endlichen liegt.

einer Hölder-Bedingung genügen sollen. Unter Wahrung dieser Forderungen denken wir uns die Koeffizienten-Matrix ein für allemal in den ganzen Raum derartig fortgesetzt, daß sie außerhalb einer hinreichend großen Kugel in die Einheitsmatrix übergeht. Die zweimal stetig differenzierbaren Lösungen einer Differentialgleichung $\Delta_2 u = 0$ mögen kurz „reguläre Potentialfunktionen" heißen, während wir die entsprechenden Lösungen der Laplaceschen Differentialgleichung $\Delta u = 0$ als „reguläre *spezielle* Potentialfunktionen" bezeichnen. Unter einem *Gebiet* verstehen wir stets eine beliebige zusammenhängende und dreidimensional offene Punktmenge, deren sämtliche Randpunkte eine beschränkte Menge bilden. Ist das Gebiet nicht beschränkt, so nehmen wir den unendlich fernen Punkt P_∞ zum Rand hinzu. Die Randwerte dürfen von irgendeiner auf der abgeschlossenen Randpunktmenge stetigen Funktion $f(p)$ [5]) gebildet werden. Für „glatte Gebiete", das heißt solche, deren Rand (nötigenfalls von P_∞ abgesehen) aus endlich vielen stetig gekrümmten Flächen besteht, ist die erste Randwertaufgabe im strengen Sinne lösbar; z. B. läßt sich die von W. Sternberg [6]) für einfach zusammenhängende glatte Gebiete entwickelte Methode auf beliebige glatte Gebiete genau so wie in der speziellen Potentialtheorie übertragen. Ehe wir zu beliebigen Gebieten übergehen, seien einige Bemerkungen über *potentialkonvexe* Funktionen vorausgeschickt, wobei wir unter einer solchen eine stetige Funktion $\varphi(P)$ verstehen, die mit einer jeden regulären Potentialfunktion $u(P)$ im Innern eines glatten Gebietes die Ungleichung $\varphi \geqq u$ erfüllt, sofern sie auf dem Rande erfüllt ist. Beispielsweise ist $\varphi(P)$ potentialkonvex, falls es dreimal stetig differenzierbar ist und der Differentialausdruck $\Delta_2 \varphi$ nicht positiv ausfällt. Diese Bemerkung führt sofort zu der Möglichkeit [7]), in jedem beschränkten abgeschlossenen Bereich B ein beliebiges Polynom $F(P)$ als Differenz zweier potentialkonvexer Funktionen φ und ψ darzustellen. Man braucht dazu nur das Maximum des Betrages von $\Delta_2 F(P)$ in diesem Bereich mit $4\pi M$ zu bezeichnen, und hat in $F(P) = \varphi(P) - \psi(P)$ die gewünschte Darstellung, falls

$$\varphi(P) = F(P) + M \iiint_B K(P, Q)\, d\tau_Q,$$

$$\psi(P) = \varphi(P) - F(P)$$

gesetzt wird; hierbei ist $K(P, Q)$ die von W. Sternberg betrachtete im Unendlichen verschwindende symmetrische Grundlösung [6]).

---

[5]) Wir werden *Randpunkte* vielfach mit kleinen Buchstaben bezeichnen, während große Buchstaben beliebige, meist innere Punkte bedeuten.

[6]) W. Sternberg, Über die linearen elliptischen Differentialgleichungen zweiter Ordnung in drei unabhängigen Veränderlichen, Math. Zeitschr. 21 (1924), S. 286—311.

[7]) Hierauf machte mich Herr W. Feller aufmerksam.

2. Existenz- und Eindeutigkeitssatz für beschränkte Gebiete.

Als Lösung im erweiterten Sinne der zum beliebigen Gebiet $\mathfrak{G}$ und der stetigen Randfunktion $f(p)$ gehörigen Randwertaufgabe suchen wir eine solche Funktion, die gleich der klassischen Lösung ist, falls diese existiert und sich sonst als Limes von gewissen klassischen Lösungen $u_i(P)$ darstellen läßt, wobei wir die $u_i(P)$ folgendermaßen definieren: wir approximieren $\mathfrak{G}$ durch eine Folge von glatten Gebieten $\mathfrak{G}_i$, die mit ihrem Rand Γ_i im Innern von $\mathfrak{G}$ liegen. Solche Folgen mögen „zulässige Approximationsfolgen" heißen. Sodann setzen wir die Funktion $f(p)$ zu einer im ganzen Raum stetigen Funktion $F(P)$ fort und lösen die durch die auf Γ_i von $F(P)$ gebildeten Randwerte gegebene Randwertaufgabe für die glatten Gebiete $\mathfrak{G}_i$. Die $u_i(P)$ sind die auf diese Weise erhaltenen in $\mathfrak{G}_i$ regulären Potentialfunktionen, die wir in $\mathfrak{G} - \mathfrak{G}_i$ durch die Gleichung $u_i(P) = F(P)$ fortgesetzt denken. Dann gilt der folgende

Existenzsatz. *Bilden die $\mathfrak{G}_i$ eine zulässige Approximationsfolge für das Gebiet $\mathfrak{G}$ und ist $F(P)$ eine stetige Fortsetzung der gegebenen Randwerte $f(p)$, so konvergieren die zugehörigen Funktionen $u_i(P)$ gegen eine in $\mathfrak{G}$ reguläre Potentialfunktion $u(P)$, die eine Lösung der zu $\mathfrak{G}$ und $f(p)$ gehörigen Randwertaufgabe heißt. Ferner konvergieren die Ableitungen erster Ordnung von $u_i(P)$ gegen die entsprechenden Ableitungen von $u(P)$. Die Konvergenz der Funktionen und der ersten Ableitungen ist gleichmäßig in jedem abgeschlossenen Teilbereich $\mathfrak{B}$ von $\mathfrak{G}$.*

Wir führen den Beweis zunächst für den Spezialfall, daß $F(P)$ potentialkonvex ist, und machen für einen Augenblick über die Folge $\mathfrak{G}_i$ die zusätzliche Voraussetzung, daß sie monoton sei, d. h. daß $\mathfrak{G}_i + \Gamma_i$ in $\mathfrak{G}_{i+1}$ enthalten ist. Wie man ohne weiteres sieht, bilden dann die Funktionen $u_i(P)$ eine abnehmende Folge, und die Behauptung folgt sofort aus dem zweiten Harnackschen Satz[8]). Es liegt auf der Hand, wie man sich von der Monotonitätsvoraussetzung über die Folge $\mathfrak{G}_i$ befreien kann, so daß die Behauptung allgemein für potentialkonvexes $F(P)$ bewiesen ist. Da aber jedes Polynom als Differenz zweier potentialkonvexer Funktionen dargestellt werden kann, ist zugleich der Fall mit erledigt, daß $F(P)$ ein Polynom ist. Wenn nunmehr $F(P)$ eine beliebige stetige Funktion ist, so approximiere man sie gleichmäßig in $\mathfrak{G} + \Gamma$ durch eine Folge von Polynomen $F_n(P)$. Die zu $F_n(P)$ und $\mathfrak{G}_i$ gehörigen Lösungen seien $u_i^{(n)}(P)$; sie konvergieren bei wachsendem i gegen die zu F_n und $\mathfrak{G}$ gehörigen Lösungen im erweiterten Sinne $u^{(n)}(P)$; diese wieder konvergieren nach dem

[8]) Vgl. W. Feller, Über die Lösungen der linearen partiellen Differentialgleichungen zweiter Ordnung vom elliptischen Typ, Math. Annalen **102** (1930), S. 633—649.

ersten Harnackschen Satz gleichmäßig gegen eine in $\mathfrak{G}$ reguläre Potentialfunktion $u(P)$. Mit Rücksicht darauf, daß die Polynome $F_n(P)$ die Funktion $F(P)$ gleichmäßig in $\mathfrak{G}$ approximieren, ergibt sich mühelos, daß auch die in der Behauptung auftretende Folge $u_i(P)$ gegen $u(P)$ konvergiert mitsamt den Ableitungen erster Ordnung, und zwar gleichmäßig im Sinne der Behauptung, womit der Existenzsatz bewiesen ist.

Der Relation $u_i(P) \to u(P)$ entnimmt man sofort die wichtige Ungleichung

$$\text{(1)} \qquad \operatorname*{Min}_{p \subset \Gamma} f(p) \leqq u(P) \leqq \operatorname*{Max}_{p \subset \Gamma} f(p),$$

wobei das Gleichheitszeichen wegen der Nichtexistenz eines Extremums im Innern des Regularitätsbereiches ausgeschlossen ist, falls $f(p)$ nicht konstant ist.

Eindeutigkeitssatz. *Die Lösung $u(P)$ hängt nur vom Gebiet $\mathfrak{G}$ und von den Randwerten $f(p)$ ab, nicht von der speziellen Wahl der Gebietsfolge $\mathfrak{G}_i$ und der Fortsetzungsfunktion $F(P)$.*

Beweis. Um die Unabhängigkeit der Lösung von der Wahl der approximierenden Gebietsfolge zu zeigen, betrachten wir zwei verschiedene zulässige Folgen $\mathfrak{G}_i$ und $\mathfrak{G}_i'$. Die Funktion $F(P)$ sei beidemal die gleiche; $u = \lim_{i \to \infty} u_i$, $u' = \lim_{l \to \infty} u_l'$ seien die beiden zugehörigen Lösungen. Wir bilden eine dritte Folge $\mathfrak{G}_k'' \to \mathfrak{G}$, die unendlich viele $\mathfrak{G}_i$ und unendlich viele $\mathfrak{G}_i'$ enthält; dieses ist stets möglich. Weil $u'' = \lim_{k \to \infty} u_k''$ existiert, ist $u'' = u$, $u'' = u'$, also $u = u'$.

Jetzt seien $F(P)$ und $F'(P)$ zwei stetige Fortsetzungen der gleichen Randfunktion $f(p)$. Wieder werde $u = \lim_{i \to \infty} u_i$, $u' = \lim_{i \to \infty} u_i'$ gesetzt, wobei beidemal die gleiche Gebietsfolge $\mathfrak{G}_i$ benutzt werde. Bezeichnen wir die Differenz $u_i(P) - u_i'(P)$ mit $w_i(P)$, so haben wir wegen $|F - F'| \leqq \varepsilon$ auf Γ_i für $i > i_\varepsilon$ und $w_i = F - F'$ auf Γ_i in $\mathfrak{G}_i + \Gamma_i$ die Ungleichung $|w_i(P)| \leqq \varepsilon$. Daraus folgt $w_i(P) \to 0$ für alle P in $\mathfrak{G}$; es ist also $u(P) = u'(P)$. Da man den allgemeinen Fall, daß gleichzeitig die Gebietsfolge $\mathfrak{G}_i$ und die Funktion $F(P)$ geändert wird, auf die beiden behandelten Spezialfälle zurückführen kann, so ist der Eindeutigkeitssatz bewiesen.

Aus ihm folgt insbesondere, daß die Lösung $u(P)$ im erweiterten Sinne mit der klassischen Lösung $U(P)$ übereinstimmt, falls diese existiert; man braucht ja nur $F(P) = U(P)$ zu setzen und hat $U(P) = u_i(P) = u(P)$.

3. Unbeschränkte Gebiete.

Ist $\mathfrak{G}$ ein beliebiges unbeschränktes Gebiet, Γ der im Endlichen gelegene Teil des Randes, $f(p)$ eine stetige auf Γ definierte Randwertfunktion und c der Wert in P_∞, so gelangen wir ähnlich wie im Fall beschränkter Gebiete auf folgende Weise zur Lösung im erweiterten Sinne: wir approxi-

mieren $\mathfrak{G}$ von innen durch eine Folge unbeschränkter glatter Gebiete $\mathfrak{G}_i$. Sodann setzen wir $f(p)$ in eine Umgebung von Γ stetig fort zu einer Funktion $F(P)$. Mit den von $F(P)$ auf Γ_i gebildeten Werten und dem Werte c im Unendlichen lösen wir durch die Funktionen $u_i(P)$ die Randwertaufgabe für $\mathfrak{G}_i$. Dann gilt wieder der

Existenz- und Eindeutigkeitssatz. *Die Folge $u_i(P)$ konvergiert gegen eine in $\mathfrak{G}$ reguläre Potentialfunktion $u(P)$. Diese Grenzfunktion hängt nur vom Gebiet $\mathfrak{G}$, den Randwerten $f(p)$ und c ab und heißt die Lösung im erweiterten Sinne der zugehörigen Randwertaufgabe. Die Konvergenz der Funktionen und ihrer ersten Ableitungen ist gleichmäßig in jedem abgeschlossenen beschränkten Teilbereich.*

Der Beweis ergibt sich fast wörtlich wie bei beschränkten Gebieten, wenn man beachtet, daß es auch dort genügt hätte, $F(P)$ nur in einer Umgebung von Γ zu definieren, und daß die Majorisierung durch die Randwerte für die Funktionen $u_i(P)$ gilt. Diese letztere Eigenschaft erlaubt, wiederum im wesentlichen alles auf monotone Folgen hinauszuspielen.

Entsprechend (1) gilt hier

$$m \leqq u(P) \leqq M, \tag{2}$$

wo m und M das Minimum und Maximum der Randwerte $f(p)$ und c ist. Ferner überzeugt man sich mühelos davon, daß $u(P)$ in P_∞ stets den vorgeschriebenen Wert c stetig annimmt.

§ 2.

Die Greensche Funktion im erweiterten Sinne.

1. Definition und einfache Eigenschaften.

Ist $K(P, Q)$ die zu $\Delta_3 u = 0$ und dem Pol Q gehörige symmetrische Grundlösung[9]), und $h(P, Q)$ die zum Gebiet $\mathfrak{G}$ und den von $K(p, Q)$ auf seinem Rand gebildeten Werten gehörige Lösung im erweiterten Sinne, so nennen wir die Funktion

$$G(P, Q) = K(P, Q) - h(P, Q) \tag{1}$$

die zum Gebiet $\mathfrak{G}$ gehörige *Greensche Funktion im erweiterten Sinne.* Ohne Schwierigkeit bestätigt man die folgenden Hilfssätze, in denen wir stets $P + Q$ voraussetzen.

a) In jedem beschränkten Gebiet gilt

$$K(P, Q) < C \frac{1}{\overline{PQ}}, \tag{2}$$

wo C unabhängig von P und Q ist.

[9]) Vgl. Anm. [6]).

b) $G(P, Q) \leqq K(P, Q)$.

c) Wenn P und Q im Innern von $\mathfrak{G}$ liegen, so gilt

$$G(P, Q) > 0. \tag{3}$$

d) Ist $\mathfrak{G}_1$ ein Teilgebiet von $\mathfrak{G}$, so ist $G_1(P, Q) \leqq G(P, Q)$ für P und Q in $\mathfrak{G}_1$.

e) Sind $\mathfrak{G}_i$ irgendwelche Gebiete, die $\mathfrak{G}$ von innen approximieren, so gilt für die zugehörigen Greenschen Funktionen im erweiterten Sinne die Relation

$$G_i(P, Q) \to G(P, Q). \tag{4}$$

Die Konvergenz ist in jedem abgeschlossenen Teilbereich gleichmäßig. (Stetige Abhängigkeit vom Gebiet.)

f) Sind Q und Q' zwei feste Punkte in $\mathfrak{G}$, während p ein Randpunkt ist, so gilt $G(P, Q') \to 0$ für $P \to p$ falls $G(P, Q) \to 0$; gilt nur für eine passende Folge $P_k \to p$ die Relation $G(P_k, Q) \to 0$, so gilt für die gleiche Folge

$$G(P_k, Q') \to 0. \tag{5}$$

Überdies gelten diese Konvergenzeigenschaften gleichmäßig für alle Q' eines jeden abgeschlossenen Teilbereiches von $\mathfrak{G}$.

(Beweis wie in der speziellen Potentialtheorie bei O. D. Kellog[10].)

2. *Zusammenhang zwischen Greenscher Funktion und Lösung der ersten Randwertaufgabe.*

$\mathfrak{G}$ sei ein beschränktes Gebiet, $f(p)$ seien die vorgegebenen Randwerte und $u(P)$ die zugehörige Lösung. Dann gilt der folgende

Satz 1. *Falls die stetige Fortsetzung der Randwerte $f(p)$ als dreimal stetig differenzierbare Funktion $F(P)$ gewählt werden kann, so gilt die Darstellung*

$$u(P) = F(P) + \frac{1}{4\pi} \iiint_{\mathfrak{G}} G(P, Q)\, \Delta_2 F(Q)\, d\tau_Q. \tag{6}$$

Hierbei ist das Integral nötigenfalls im Lebesgueschen Sinne zu nehmen.

Beweis. Zunächst ist $\Delta_2 u = 0$. Daß ferner $u(P)$ im Falle eines glatten Gebietes die Randwerte $f(p)$ annimmt, folgt sofort aus den Hilfssätzen a), b) und f). Daß die Integraldarstellung aber auch bei beliebigem Gebiet gilt, erkennt man wie in der speziellen Potentialtheorie[10], indem man $\mathfrak{G}$ durch glatte Gebiete $\mathfrak{G}_i$ und $u(P)$ durch die zu $F(P)$ und $\mathfrak{G}_i$ gehörigen Lösungen approximiert. Wendet man nämlich auf diese Lösungen die Dar-

[10]) Vgl. Anm. [3]).

stellung (6) an, so ergibt sich die Behauptung unter Benutzung der Hilfssätze a), b) und e).

Gestattet jetzt $f(p)$ keine dreimal stetig differenzierbare Funktion als stetige Fortsetzung $F(P)$, so bezeichne $F_n(P)$ eine Polynomfolge, die $F(P)$ in $\mathfrak{G}+\Gamma$ gleichmäßig approximiert. Für die Lösung $u^{(n)}(P)$, die zu den Randwerten $F_n(P)$ gehört, gilt die Darstellung (6), indem man u und F durch $u^{(n)}$ und F_n ersetzt. Nun ist aber $u = \lim_{n\to\infty} u^{(n)}$. Mithin gilt unter Benutzung von $F_n(P) \to F(P)$

$$(7) \qquad u(P) = F(P) + \lim_{n\to\infty} \frac{1}{4\pi} \iiint_{\mathfrak{G}} G(P,Q)\, \Delta_2 F_n(Q)\, d\tau_Q .$$

Mit Rücksicht auf den Hilfssatz e) schließt man nunmehr leicht aus (6) mit Hilfe einfacher Überlegungen auf den folgenden

Satz 2. *Sind $\mathfrak{G}_i$ beliebige, nicht notwendig glatte Gebiete, die das beschränkte Gebiet $\mathfrak{G}$ von innen approximieren, und sind $u_i(P)$ und $u(P)$ die zugehörigen Lösungen im erweiterten Sinne, dann besteht die Relation $u_i(P) \to u(P)$. In jedem abgeschlossenen Teilbereich ist die Konvergenz gleichmäßig.*

Als weitere Anwendung der Integraldarstellung ergibt sich unter Benutzung der Hilfssätze a), b) und f) ähnlich wie in der speziellen Potentialtheorie [10]) der für das Studium der Randwertannahme grundlegende

Satz 3. *Wenn die Greensche Funktion $G(P,Q)$ für einen festen Punkt Q aus $\mathfrak{G}$ gegen Null konvergiert, falls sich P dem Randpunkt p auf einer bestimmten Folge P_k nähert, so strebt $u(P_k)$ gegen den Randwert $f(p)$; nimmt also speziell die Greensche Funktion in p den Wert Null stetig an, so nimmt $u(P)$ den Wert $f(p)$ in p stetig an.*

Durch diesen Satz ist die Frage der Randwertannahme der Lösung $u(P)$ auf die entsprechende Frage bei der Greenschen Funktion zurückgeführt. Wir werden in dieser Richtung noch einen Schritt weitergehen und führen zu diesem Zwecke im nächsten Paragraphen zwei Hilfsbegriffe ein.

§ 3.

Das normierte Potential und die Kapazität einer beliebigen beschränkten Punktmenge.

1. Definitionen und einfache Eigenschaften.

Es liege eine beliebige beschränkte Punktmenge B vor, B' sei ihre abgeschlossene Hülle. Die Komplementärmenge von B' enthält ein Gebiet Ω, das sich ins Unendliche erstreckt. Sein Rand Σ ist in der Randmenge von B enthalten. Indem wir einen in der speziellen Potentialtheorie neuerdings benutzten Begriff verallgemeinern, verstehen wir unter dem *„normierten Po-*

tential" $v(P)$ von B die Lösung im erweiterten Sinne der ersten Randwertaufgabe für das Gebiet Ω mit den Randwerten Eins auf Σ und Null im Unendlichen. Ist $\mathfrak{F}$ irgendeine glatte geschlossene Fläche ohne Doppelpunkte, die B' umschließt, n die gegen B' weisende Flächennormale, und wird

$$\sum_{i,k=1}^{3} a_{i,k} \cos(x_i, n) \frac{\partial}{\partial x_k} = \frac{\delta}{\delta l} \tag{1}$$

gesetzt, so definieren wir die *Kapazität* c von B durch das von $\mathfrak{F}$ unabhängige Integral

$$c = \frac{1}{4\pi} \int\int_{\mathfrak{F}} \frac{\delta}{\delta l} v(Q)\, do. \tag{2}$$

Wie man sieht, hängt $v(P)$ und c nur von dem Teil Σ der Randmenge von B ab, der zugleich Rand von Ω ist. Speziell ist $c(B) = c(B') = c(\Sigma)$. Im übrigen hängt c noch von der Wahl des Differentialausdruckes $\Delta_2 u$ ab.

Falls Ω glatt ist, kann man für $v(P)$ und c explizite Darstellungen geben. Zunächst ist wie in der speziellen Potentialtheorie

$$v(P) = \int\int_{\Sigma} \varrho(p) K(p, P)\, do_p, \tag{3}$$

wo sich die Belegung $\varrho(p)$ als nicht-triviale Lösung einer homogenen Integralgleichung gewinnen läßt. Da, wie man sich leicht überzeugt, die von W. Sternberg[6]) bewiesene Sprungrelation für Potentialfunktionen der Form (3) gleichmäßig in p gelten, erhält man für glattes Ω aus (2) und (3) für die Kapazität die Darstellung

$$c = \int\int_{\Sigma} \varrho(p)\, do. \tag{4}$$

Da ferner, wie man ebenfalls bestätigt, die benutzte Sprungrelation auch bei Randannäherung längs der durch $\frac{\delta}{\delta l}$ bestimmten Richtung gültig bleibt, stellt man leicht mit Rücksicht darauf, daß $v(P) < 1$ in Ω ist, fest, daß $\varrho(p) \geqq 0$ gilt. Demnach erhält man, falls Ω glatt ist, aus (3) und (4) die die für später wichtige Ungleichung

$$c \cdot \operatorname*{Min}_{p \subset \Gamma} K(p, P) \leqq v(P) \leqq c \cdot \operatorname*{Max}_{p \subset \Gamma} K(p, P). \tag{5}$$

Um diese Ungleichung auszudehnen auf den Fall einer beliebigen beschränkten Punktmenge B mit dem Potential $v(P)$, schließen wir B durch glatte Flächen Σ_i ein; das außerhalb Σ_i liegende unendliche Gebiet heiße Ω_i; das zugehörige normierte Potential $v_i(P)$. Ist Σ_i eine solche Folge, daß die zugehörigen Ω_i das Gebiet Ω ausschöpfen, und zwar in monotoner Weise, so sind die $v_i(P)$ monoton bezüglich i und konvergieren nach § 1 gegen $v(P)$.

Da auch die ersten Ableitungen in jedem abgeschlossenen Teilbereich gleichmäßig gegen die zugehörigen Ableitungen konvergieren, so haben wir wegen (2) sofort die Relation

$$c_i \to c, \tag{6}$$

wo c_i die Kapazität von Σ_i ist. Nunmehr ergibt sich ohne weiteres die Allgemeingültigkeit von (5).

Wie man dieser Ungleichung entnimmt, folgt $v(P) \equiv 0$ aus $c = 0$ und umgekehrt.

Ferner erwähnen wir noch eine weitere wichtige Ungleichung: Sind B und E zwei Punktmengen, so gilt

$$c(B) \leqq c(B+E) \leqq c(B) + c(E). \tag{7}$$

Man beweist diese Ungleichung ähnlich wie in der speziellen Potentialtheorie [3]) [3a]) unter Benutzung von (2).

2. Vergleich der Kapazitäten einer Punktmenge bezüglich verschiedener Differentialausdrücke.

Zunächst geben wir für das normierte Potential und die Kapazität einer Punktmenge B mit glattem Ω eine Charakterisierung durch ein Minimumproblem [11]). Wir behaupten, daß das Potential $v(P)$ die Lösung des folgenden Variationsproblems ist:

$$\mathfrak{D}[\varphi] = \iiint_{\Omega} \left\{ \sum_{i,k=1}^{3} a_{i,k} \frac{\partial \varphi}{\partial x_i} \frac{\partial \varphi}{\partial x_k} \right\} d\tau = \text{Min}, \tag{8}$$

und daß dieses Minimum $\mathfrak{D}[v]$ die Gleichung erfüllt

$$\mathfrak{D}[v] = 4\pi c. \tag{9}$$

Hierbei sind alle Funktionen φ zulässig, die in Ω stetig sind, auf Σ den Wert 1, in P_∞ den Wert 0 annehmen; die ersten Ableitungen von φ sollen in Ω stückweise stetig sein und dem Ausdruck $\mathfrak{D}[\varphi]$ einen endlichen Wert erteilen.

Man überzeugt sich hiervon, indem man Σ durch Parallelflächen im Abstande h approximiert und auf die so entstehenden Teilgebiete Ω_h von Ω die Greensche Formel anwendet. Läßt man h gegen Null streben und beachtet, daß dabei $\frac{\delta v}{\delta l}$ gleichmäßig gegen $4\pi\varrho(p)$ konvergiert, so ergibt sich die Behauptung.

Nunmehr sind wir in der Lage, eine Ungleichung herzuleiten, die die Grundlage dafür bildet, daß die Einteilung der Randpunkte eines Gebietes

---

[11]) Vgl. G. Bouligand, Ensembles impropres et nombre dimensionel, Bull. des Sciences math. **52** (12. sér.) (1928), S. 320—344 und S. 361—376. Insbesondere S. 373.

in drei Klassen gemäß dem Randwertverhalten der Lösung im erweiterten Sinne unabhängig ist von der speziellen jeweiligen Differentialgleichung. Die fragliche Ungleichung lautet

$$m c_{\Delta} \leqq c_{\Delta_2} \leqq M c_{\Delta}, \tag{10}$$

wo c_{Δ} und c_{Δ_2} die zu den Differentialausdrücken Δu und $\Delta_2 u$ gehörigen Kapazitäten einer Punktmenge B sind, während m und M positive Konstanten bezeichnen, die nur vom Differentialausdruck $\Delta_2 u$, nicht aber von B abhängen.

Beim Beweise genügt es, sich auf glatte B zu beschränken. Setzen wir zur Abkürzung

$$D[\varphi] = \iiint_{\Omega} \{\varphi_x^2 + \varphi_y^2 + \varphi_z^2\}\, dx\, dy\, dz,$$

wo x, y, z statt x_1, x_2, x_3 geschrieben ist, so folgt mit Rücksicht darauf, daß die $a_{i,k}(P)$ des elliptischen Differentialausdruckes $\Delta_2 u$ im ganzen Raum stetig und außerhalb einer gewissen Kugel konstant sind, die Ungleichung

$$m D[\varphi] \leqq \mathfrak{D}[\varphi] \leqq M D[\varphi]$$

für alle φ, die für das Variationsproblem (8) zulässig sind; hierbei sind m und M positive von φ und Ω unabhängige Konstanten. Ferner ist nach (9)

$$4\pi c_{\Delta} = D[w], \qquad 4\pi c_{\Delta_2} = \mathfrak{D}[v],$$

wo w und v die zu B gehörigen normierten Potentiale bezüglich Δu und $\Delta_2 u$ sind. Wegen $\mathfrak{D}[v] \leqq \mathfrak{D}[w]$ ist $\mathfrak{D}[v] \leqq M D[w]$, also auch $c_{\Delta_2} \leqq M c_1$; ähnlich folgt $m c_{\Delta} \leqq c_{\Delta_2}$. Damit ist (10) bewiesen. Entsprechende Ungleichungen gelten natürlich für die Kapazitäten bezüglich zweier beliebiger Differentialausdrücke $\Delta_2 u$ und $\Delta_2^* u$.

§ 4.

Das Verhalten der Lösung am kapazitätslosen Bestandteil des Randes.

Wir sagen, daß die Menge B in einem ihrer Punkte P die *lokale Kapazität Null* besitzt, wenn es eine hinreichend kleine Kugel K um P gibt, derart, daß die in und auf ihr liegende Teilmenge von B die Kapazität Null besitzt. Jede abgeschlossene Menge von solchen Punkten verschwindender lokaler Kapazität hat mit Rücksicht auf § 3 (7) nach dem Heine-Borelschen Satze gleichfalls die Kapazität Null. Ferner nennen wir eine Menge kapazitätslos, wenn jede abgeschlossene Teilmenge von ihr die Kapazität Null besitzt. Insbesondere ist die Gesamtheit der Punkte von B, in denen B die lokale Kapazität Null besitzt, eine solche. Sie heiße der *kapazitätslose Bestandteil* von B und werde mit Λ bezeichnet. Nach § 3 (10) hängt Λ nicht von der speziellen Wahl der Differentialgleichung $\Delta_2 u$ ab.

Über das Verhalten der Lösung im erweiterten Sinne am kapazitätslosen Bestandteil des Randes gelten die gleichen Sätze, wie sie O. D. Kellog und F. Vasilesco[3]) gefunden haben. Da sich auch die von dem letzteren benutzten Beweise ohne Schwierigkeit übertragen lassen, genügt es, das Resultat anzugeben:

$\mathfrak{G}$ sei irgendein Gebiet, γ ein Teil seines Randes mit der Eigenschaft, daß $\mathfrak{G} + \gamma$ wieder ein Gebiet ist. Falls γ kapazitätslos ist, so kann jede Potentialfunktion $u(P)$, die in der Umgebung von γ beschränkt und regulär ist, auf γ so definiert werden, daß sie in γ regulär wird. Umgekehrt, falls dies stets möglich ist, so ist γ kapazitätslos. Insbesondere erfüllt der kapazitätslose Bestandteil Λ des Randes die über γ gemachten Voraussetzungen.

Die Punktmenge $\mathfrak{G}' = \mathfrak{G} + \Lambda$, die hiernach wieder ein Gebiet ist, heißt *reduziert*. Ihr Rand $\Gamma - \Lambda$ enthält keinen kapazitätslosen Bestandteil mehr und heißt gleichfalls reduziert.

§ 5.

Das Verhalten der Lösung in einem beliebigen Randpunkt.

1. Gedankengang der folgenden Untersuchung.

Die Methode, mit der wir das Verhalten der Lösung in einem Punkte des reduzierten Randes untersuchen, besteht darin, daß wir diese Frage zurückführen auf das Verhalten der Greenschen Funktion sowie der normierten Potentiale gewisser Punktmengen, die nur von der Umgebung des betrachteten Randpunktes abhängen. Dadurch werden wir zu einem notwendigen und hinreichenden Kriterium für die Regularität[12]) eines Randpunktes geführt, das die Unabhängigkeit der in der Einleitung ausgesprochenen Einteilung der Randpunkte in drei Klassen von der speziellen Wahl der Differentialgleichung $\Delta_2 u = 0$ in Evidenz setzt, womit der Anschluß an die für die speziellen Potentialfunktionen entwickelte Theorie gewonnen ist.

2. Die Zurückführung auf gewisse spezielle Randwertaufgaben.

Aus der Definition der Greenschen Funktion im erweiterten Sinne, sowie aus Satz 3 § 2 ergibt sich sofort der folgende

Satz 1. *Notwendig und hinreichend für die Regularität des Randpunktes p ist es, daß die Greensche Funktion des Gebietes $\mathfrak{G}$ für irgendeinen Aufpunkt Q in p den Wert Null stetig annimmt, daß also die folgende Relation gilt*

$$G(P, Q) \to 0 \quad \text{für} \quad P \to p. \tag{1}$$

[12]) Definition in der Einleitung.

Satz 2. *Sind $\mathfrak{G}$ und $\mathfrak{H}$ zwei Gebiete mit dem gemeinsamen Randpunkt p, und ist $\mathfrak{G} < \mathfrak{H}$, so ist der Randpunkt p bezüglich $\mathfrak{G}$ regulär, falls er es bezüglich $\mathfrak{H}$ ist.*

Sind nämlich $G(P, Q)$ und $H(P, Q)$ die zu $\mathfrak{G}$ und $\mathfrak{H}$ gehörigen Greenschen Funktionen im erweiterten Sinne, so gilt $0 < G(P, Q) \leqq H(P, Q)$ für P und Q in $\mathfrak{G}$, also $G(P, Q) \to 0$ für $P \to p$, falls das Entsprechende für $H(P, Q)$ gilt.

Um zu zeigen, daß für die Regularität von p ebenso wie das Verhalten der Greenschen Funktion auch das Verhalten gewisser normierter Potentiale maßgebend ist, brauchen wir eine Ungleichung zwischen dem normierten Potential $v(P)$ einer beliebigen Punktmenge B und der Greenschen Funktion $H(P, Q)$ des zu B gehörigen Außengebietes Ω. Es sei Q ein Punkt aus Ω, K eine beliebige in Ω enthaltene Kugel um Q, auf welcher das Maximum von $v(P)$ mit m, das Maximum von $H(P, Q)$ mit M bezeichnet werde $(0 \leqq m < 1)$. Dann gilt in dem ganzen außerhalb K liegenden Teile von Ω die Ungleichung

$$H(P, Q) \leqq \frac{M}{1-m}(1 - v(P)). \tag{2}$$

Man beweist sie genau so wie die entsprechende für spezielle Potentialfunktionen, die O. D. Kellog [3]) [a]) angegeben hat. Nunmehr sei $\mathfrak{K}_\alpha$ die Kugel um p mit dem Radius α; die Menge derjenigen in ihr enthaltenen Punkte, die nicht zu $\mathfrak{G}$ gehören, heiße E_α, während $v_\alpha(P)$ das normierte Potential von E_α sei. Sobald α hinreichend klein ist, gehört $\mathfrak{G}$ dem unendlichen Außengebiet Ω_α von E_α an und p ist ein Punkt seines Randes Σ_α. Dann gilt folgender

Satz 3. *Für die Regularität von p ist es notwendig und hinreichend, daß für einen genügend kleinen Wert α das Potential $v_\alpha(P)$ den richtigen Randwert annimmt, d. h. daß $v_\alpha(P) \to 1$ für $P \to p$ ist.*

Ist nämlich $v_\alpha(P) \to 1$, so gilt nach (2) für die Greensche Funktion $H(P, Q)$ von Ω_α die Beziehung $H(P, Q) \to 0$ für $P \to p$, so daß p nach Satz 1 bezüglich Ω_α, also nach Satz 2 auch bezüglich $\mathfrak{G}$ regulär ist. Die Umkehrung wird genau so bewiesen wie in der speziellen Potentialtheorie [13]).

Bei diesem Kriterium kommt nur eine beliebig kleine Umgebung $\mathfrak{G}_\alpha$ des Gebietes $\mathfrak{G}$ ins Spiel, wodurch der *rein lokale Charakter der Regularität* eines Randpunktes in Evidenz gesetzt wird. Sagen wir kurz, ein Gebiet $\mathfrak{G}$ sei in einer Umgebung von p Teilgebiet eines Gebietes $\mathfrak{H}$, falls für ein passendes $\alpha > 0$ die in einer Kugel $\mathfrak{K}_\alpha$ enthaltene Teilmenge $\mathfrak{G}_\alpha$ von $\mathfrak{G}$ Teilmenge der entsprechenden Menge $\mathfrak{H}_\alpha$ ist, so ergibt sich aus dem lokalen Charakter der Regularität und dem Satz 2 sofort das folgende

[13]) Vgl. N. Wiener unter Anm. [2])[b]), S. 135.

Majorantenprinzip. *Ist p gemeinsamer Randpunkt der beiden Gebiete $\mathfrak{G}$ und $\mathfrak{H}$, ist ferner $\mathfrak{G}$ in einer Umgebung von p Teilmenge von $\mathfrak{H}$, so ist der Randpunkt p für $\mathfrak{G}$ regulär, falls er es für $\mathfrak{H}$ ist, dagegen singulär für $\mathfrak{H}$, falls er es für $\mathfrak{G}$ ist.*

3. *Das Wienersche Reihenkriterium und seine Übertragung auf die Differentialgleichungen der Form $\Delta_2 u = 0$.*

Nach Ungleichung (5) des § 3 besteht zwischen dem Potential einer Punktmenge und ihrer Kapazität ein Zusammenhang, der es erwarten läßt, daß sich auch mit Hilfe der Kapazitäten der Mengen E_α ein Kriterium für die Regularität eines Randpunktes finden läßt. Ein solches notwendiges und hinreichendes Kriterium wurde in der Tat für die spezielle Potentialtheorie von N. Wiener[2)b)] hergeleitet; man kann es nach einer Bemerkung von O. D. Kellog und F. Vasilesco[14)] in folgender Form aussprechen: λ sei eine positive Zahl mit $0 < \lambda < 1$; k_i die Menge der nicht zu $\mathfrak{G}$ gehörigen Punkte in und auf der Kugel $\mathfrak{K}_{\lambda^i}$ um p mit dem Radius λ^i; (k_i) die Kapazität von k_i genommen in bezug auf die Laplacesche Differentialgleichung $\Delta u = 0$. Der Randpunkt p ist dann regulär oder singulär bezüglich $\Delta u = 0$, je nachdem die Reihe

$$S = \sum_{i=1}^{\infty} \frac{(k_i)}{\lambda^i}, \qquad (S = S(p)), \tag{3}$$

für irgendein zulässiges λ divergiert oder konvergiert.

Ehe wir zeigen, daß dies Kriterium auch für die Differentialgleichungen $\Delta_2 u = 0$ gilt, schicken wir einige Bemerkungen voraus, die sich in der Hauptsache auf das Konvergenzverhalten von S und ähnlichen Reihen beziehen.

a) Bezeichnen wir mit $(k_i)_{\Delta_2}$ die Kapazität von k_i bezüglich $\Delta_2 u = 0$ und mit $S_2 = S_2(p)$ die Reihe

$$S_2 = \sum_{i=1}^{\infty} \frac{(k_i)_{\Delta_2}}{\lambda^i}, \tag{4}$$

so haben nach § 3 (10) die Reihen S_2 und S das gleiche Konvergenzverhalten. Wir schreiben dafür kurz

$$S_2 \sim S. \tag{5}$$

Hiernach können wir bei Konvergenzbetrachtungen einfach (k_i) statt $(k_i)_{\Delta_2}$ schreiben.

[14)] O. D. Kellog und V. Vasilesco, A Contribution to the Theory of Capacity, Journ. of Math. **51** (1929), S. 515—526, insbes. S. 518.

b)

$$(6) \qquad \sum_{i=1}^{\infty} \frac{(k_i - k_{i+1})_{\Delta_2}}{\lambda^i} \sim \sum_{i=1}^{\infty} \frac{(k_i)_{\Delta_2}}{\lambda^i}.$$

(Zu dem einfachen Beweis vergleiche Anm. [14]).)

c) Der Bereich B möge durch eine nicht-singuläre lineare Transformation $P = \mathfrak{T}\,\bar{P}$ aus einem Bereich $\bar{B}$ hervorgehen, kurz $B = \mathfrak{T}\,\bar{B}$, wobei der Differentialausdruck $\bar{\Delta}_2\,\bar{u}(\bar{P})$ in $\Delta_2 u(P)$ übergehen möge. Dann gilt für die Kapazitäten von B und $\bar{B}$

$$(7) \qquad (B)_{\Delta_2} = |\mathfrak{T}|\,(\bar{B})_{\bar{\Delta}_2},$$

wo $|\mathfrak{T}|$ die Determinante von $\mathfrak{T}$ ist. Für glatte B sieht man das sofort aus der Darstellung (9) § 3:

$$(8) \qquad 4\pi(B)_{\Delta_2} = \mathfrak{D}_{\Omega}[v], \qquad 4\pi(\bar{B})_{\bar{\Delta}_2} = \overline{\mathfrak{D}}_{\bar{\Omega}}[\bar{v}].$$

Da nämlich durch die Transformation das auf $\Delta_2 u$ bezügliche normierte Potential $v(P)$ von B in das auf $\bar{\Delta}_2\,\bar{u}$ bezügliche Potential $\bar{v}(\bar{P})$ von $\bar{B}$ übergeht, folgt aus (8) für glatte B sofort wegen

$$\mathfrak{D}_{\Omega}[v] = |\mathfrak{T}|\,\overline{\mathfrak{D}}_{\bar{\Omega}}[\bar{v}]$$

die behauptete Gleichung (7). Durch eine Approximationsbetrachtung ergibt sich ihre Richtigkeit für beliebige Punktmengen B.

Aus (7) und den vorhergehenden Bemerkungen folgt, daß sich das Konvergenzverhalten von S durch eine lineare Transformation nicht ändert, (λ wird dabei festgehalten). Kurz:

$$(9) \qquad S_2 \sim \bar{S}_2.$$

d) Ferner machen wir eine Feststellung bezüglich der symmetrischen Grundlösung $K(P, Q)$ in der Umgebung eines solchen Punktes p, in dem die Koeffizientenmatrix $a_{i,k}(P)$, die zum Differentialausdruck $\Delta_2 u$ gehört, gleich der Einheitsmatrix ist, ohne daß sie in der Umgebung von p konstant zu sein braucht. Wir sagen dann, $\Delta_2 u = 0$ besitzt in P die *Normalform*. Schreiben wir die Grundlösung in der Form [15])

$$(10) \qquad K(P, Q) = \frac{1}{E(Q)} \left\{ \frac{1}{\sqrt{\sum_{i,k=1}^{3} A_{i,k}(P)\,(\xi_i - x_i)\,(\xi_k - x_k)}} + W(P, Q) \right\},$$

wobei sich x_i auf P, ξ_i auf Q bezieht und $A_{i,k}(P)$ die durch die Determinante $|a_{i,k}(P)|$ dividierte und mit Vorzeichen versehene Unterdeterminante von $a_{i,k}(P)$ ist, wo ferner $E(Q)$ eine stetige positive Funktion ist mit $E(p) = 1$ und $\overline{PQ}\,W(P, Q) \to 0$ für $\overline{PQ} \to 0$, gleichmäßig für

[15]) Vgl. Anm. [6]).

alle voneinander verschiedenen Punkte P und Q aus $\mathfrak{G}+\Gamma$, so ergibt sich in unserm Fall, daß es zu jedem $\varepsilon>0$ ein $\alpha_\varepsilon>0$ gibt, derart, daß die Ungleichung

$$|r\,K(P,Q)-1|<\varepsilon \quad \text{mit} \quad r=\overline{PQ} \tag{11}$$

für $\overline{Pp}<\alpha_\varepsilon$, $\overline{Qp}<\alpha_\varepsilon$, d. h. in $\mathfrak{R}_{\alpha_\varepsilon}$ gilt. In $\mathfrak{R}_{\alpha_\varepsilon}$ ist also

$$\frac{1-\varepsilon}{K(P,Q)}\leqq\overline{PQ}\leqq\frac{1+\varepsilon}{K(P,Q)}. \tag{12}$$

e) Schließlich erinnern wir an die Ungleichung (5) § 3. Sie nimmt, wenn P und die Punktmenge B, um deren Kapazität und Potentiale es sich handelt, in $\mathfrak{R}_{\alpha_\varepsilon}$ liegt, wegen (12) die Gestalt an

$$(1-\varepsilon)\,c\cdot\operatorname*{Min}_{Q\subset\Gamma}\frac{1}{\overline{PQ}}\leqq v(P)\leqq(1+\varepsilon)\,c\cdot\operatorname*{Max}_{Q\subset\Gamma}\frac{1}{\overline{PQ}}. \tag{13}$$

Nunmehr kommen wir zu dem Reihenkriterium für die Differentialgleichung $\Delta_2 u=0$. Es lautet:

Der Randpunkt p ist regulär oder singulär in bezug auf die Differentialgleichung $\Delta_2 u=0$, je nachdem die Reihe (3) *für irgendein λ mit $0<\lambda<1$ divergiert oder konvergiert, p ist also regulär für alle Differentialgleichungen $\Delta_2 u=0$ oder für keine.*

Nach den Vorbemerkungen können wir uns beim Beweis sehr kurz fassen. S; also auch S_2, sei divergent. Zunächst ist es klar, daß p bezüglich $\Delta_2 u$ und $\mathfrak{G}$ regulär ist oder nicht, je nachdem $\bar{p}$ bezüglich $\overline{\Delta_2 u}$ und $\overline{\mathfrak{G}}$ es ist, wobei die überstrichenen Symbole mit den unüberstrichenen durch eine lineare Transformation $P=\mathfrak{T}\bar{P}$ verbunden sind. Da sich nach (9) auch das Konvergenzverhalten der Reihe (4) durch eine solche Transformation nicht ändert, dürfen wir annehmen, daß $\Delta_2 u=0$ im Punkte p die Normalform besitzt. Nach der Bemerkung b) ist mit (4) auch die Reihe

$$S_2^*=\sum_{i=1}^{\infty}\frac{(k_i-k_{i+1})_{\Delta_2}}{\lambda^i} \tag{16}$$

divergent. Nunmehr ist zum Beweis der Regularität von p nach Satz 3 nur zu zeigen, daß für einen geeigneten Wert α das Potential $v_\alpha(P)$ in p den Wert 1 annimmt. Dies läßt sich aber unter Benutzung von (13) aus der Divergenz von S_2^* ebenso beweisen, wie es O. D. Kellog in der speziellen Potentialtheorie getan hat[16]).

Jetzt sei (3), also auch S_2^* konvergent. Nach c) dürfen wir annehmen, daß $\Delta_2 u=0$ in p die Normalform besitzt. Wir bestimmen dann eine positive ganze Zahl m derart, daß erstens $\lambda^m<\alpha_\varepsilon$ wird, wo α_ε der Unglei-

[16]) Vgl. O. D. Kellog unter Anm. [8])[b]), S. 332 ff.

chung (11) gemäß mit $\varepsilon = 1$ bestimmt ist, und daß zweitens

$$\sum_{i=m}^{\infty} \frac{\gamma_i}{\lambda^i} < \frac{\lambda}{8},$$

wo $\gamma_i = (k_i - k_{i+1})_{\Delta_2}$ ist, wird. Zum Beweis der Singularität von p genügt es zu zeigen, daß das Potential $v_\alpha(P)$ der Menge E_α, $\alpha = \lambda^m$, im Punkte p den Wert 1 nicht stetig annimmt. Dies aber stellt man wieder unter Benutzung von (7) § 3 und (13) genau durch die gleichen Überlegungen fest, mit denen O. D. Kellog die entsprechende Tatsache in der speziellen Potentialtheorie bewiesen hat[16]).

4. Endgültige Ergebnisse.

Wir vervollständigen die bisher gewonnenen Resultate durch einige Aussagen, die sich auf die singulären Randpunkte beziehen. Zunächst gilt für die singulären Punkte des reduzierten Randes folgender

Satz 4. *Jeder singuläre Punkt p des reduzierten Randes läßt sich durch eine solche Punktfolge P_k aus $\mathfrak{G}$ approximieren, daß $u(P_k) \to f(p)$ gilt.*

Der Beweis erledigt sich durch die Bemerkung, daß er genau der gleiche ist wie der von F. Vasilesco[3]) für den entsprechenden Satz bei harmonischen Funktionen benutzte.

Die weiteren Aussagen beziehen sich auf den kapazitätslosen Bestandteil des Randes.

Satz 5. *Die Punkte p des kapazitätslosen Bestandteils Λ des Randes sind singulär und dadurch charakterisiert, daß die zugehörigen Reihen $S(p)$ abbrechen.*

Die Singularität von p folgt sogleich aus Satz 3. Daß $S(p)$ abbricht, ergibt sich sofort aus dem Satz von § 4. Umgekehrt ist klar, daß ein Randpunkt, für den $S(p)$ abbricht, die lokale Kapazität Null besitzt, also zu Λ gehört.

Ehe wir zu dem nächsten auf Λ bezüglichen Satz übergehen, ist eine Bemerkung über kubische Gebiete, d. h. solche, deren Rand aus endlich vielen achsenparallelen Ebenenstücken besteht, einzuschieben. Offenbar kann man mit kubischen Gebieten jedes beliebige Gebiet $\mathfrak{G}$ im Sinn von § 1 approximieren. Man weiß ferner, daß die sämtlichen Punkte eines kubischen Gebietes für die spezielle Potentialgleichung regulär sind[17]). Mithin ist auch die Randwertaufgabe der Differentialgleichung $\Delta_2 u = 0$ für kubische Gebiete stets im strengen Sinn lösbar. Daher können wir an Stelle der in § 1 zur Konstruktion der Lösung im erweiterten Sinne benutzten Approximation durch glatte Gebiete eine solche durch kubische Gebiete verwenden.

[17]) Vgl. z. B. N. Wiener. loc. cit. [2])[b]).

Mit Hilfe dieser Bemerkung läßt sich der folgende Eindeutigkeitssatz ganz ähnlich beweisen, wie es F. Vasilesco mit dem entsprechenden Satz in der speziellen Potentialtheorie getan hat[3]). Es genügt daher das Ergebnis anzugeben.

Satz 6. *Zwei Lösungen im erweiterten Sinne der ersten Randwertaufgabe sind identisch, wenn nur ihre Randwertvorgaben auf dem reduzierten Teil des Randes übereinstimmen.*

Wendet man die Integraldarstellung von § 2 auf die zu $\mathfrak{G}$ und seinem reduzierten Gebiet $\mathfrak{G}'$ gehörigen Lösungen $u(P)$ und $u'(P)$ bei auf $\Gamma - \Lambda$ gleichen Randwertvorgaben an, so ergibt sich sofort der folgende

Satz 7. *Der kapazitätslose Bestandteil des Randes besitzt stets das dreidimensionale Maß Null.*

Abschließend erkennt man, daß für die Differentialgleichungen der Form $\Delta_2 u = 0$ die Resultate dieses und des vorangehenden Paragraphen das in der Einleitung zusammengefaßte Hauptergebnis in sich enthalten.

§ 6.

Die allgemeine selbstadjungierte elliptische Differentialgleichung $\Delta_2 u + a u + b = 0$.

$\mathfrak{G}$ sei ein beliebiges beschränktes Gebiet; die Funktionen $a(P)$ und $b(P)$ seien mit ihren Ableitungen erster Ordnung in $\mathfrak{G}$ stetig. Um die im vorangehenden für die Lösungen von $\Delta_2 u = 0$ entwickelte Theorie auf den allgemeinen Fall zu übertragen, bezeichnen wir mit $w(P)$ die zu $\mathfrak{G}$ und den stetigen Randwerten $f(p)$ gehörige Lösung im erweiterten Sinne von $\Delta_2 w = 0$, setzen zur Abkürzung $a w + b = h$ und machen für $u(P)$ den Ansatz $u = v + w$, der für $v(P)$ die Differentialgleichung

$$\Delta_2 v + a v + h = 0 \tag{1}$$

mit den Randwerten Null ergibt. Setzen wir noch

$$H(P) = \frac{1}{4\pi} \iiint_{\mathfrak{G}} G(P, Q)\, h(Q)\, d\tau_Q, \tag{2}$$

wo $G(P, Q)$ die zum Gebiet $\mathfrak{G}$ gehörige Greensche Funktion im erweiterten Sinne bezüglich der Differentialgleichung $\Delta_2 w = 0$ ist, so definieren wir allgemein $v(P)$ durch die Integralgleichung

$$v(P) = \frac{1}{4\pi} \iiint_{\mathfrak{G}} G(P, Q)\, a(Q)\, v(Q)\, d\tau_Q + H(P) \tag{3}$$

bzw., falls $a(P) \equiv 0$, durch

$$v(P) = H(P). \tag{4}$$

Auf die Integralgleichung (3) läßt sich die übliche Theorie anwenden; insbesondere folgt aus der bekannten Alternative, daß entweder das zu den Randwerten Null gehörige Eigenwertproblem

$$\Delta_2 v + \lambda a(P) v = 0 \tag{5}$$

mit dem Eigenwert $\lambda = 1$ oder aber die Differentialgleichung

$$\Delta_2 u + a(P) u + b(P) = 0 \tag{6}$$

mit beliebigen Randwerten und beliebiger Funktion $b(P)$ eine Lösung im erweiterten Sinne besitzt. Ferner folgt die Existenz von abzählbar vielen, sich im Endlichen nicht häufenden Eigenwerten mit je endlich vielen Eigenfunktionen im erweiterten Sinne. Die Frage nach der Annahme der Randwerte für die Lösung von (6) und für die Eigenfunktionen läßt sich sehr einfach erledigen. Wir behaupten:

Die durch das Konvergenzverhalten der in § 5 *eingeführten Reihe* $S(p)$ *bestimmte Einteilung eines beliebigen beschränkten Gebietes in drei Klassen besitzt auch für die Lösungen im erweiterten Sinne der Randwertaufgabe der Differentialgleichung* (6) *wie auch für die Eigenfunktionen von* (5) *die in der Einleitung ausgesprochene Bedeutung.*

Wendet man nämlich die Ergebnisse von § 4 und § 5 auf $w(P)$ und $H(P)$, auf letzteres ferner die Überlegungen von § 2 an, so erkennt man zunächst, daß die fragliche Klasseneinteilung für die Funktionen $w(P)$ und $H(P)$ gültig ist. Also wegen $u(P) = v(P) + w(P)$ und (4) auch für $u(P)$, falls $a(P) = 0$. Sie bleibt aber auch bei nicht identisch verschwindenden $a(P)$ gültig, wie man ohne Schwierigkeit erkennt, wenn man die für die Greensche Funktion im erweiterten Sinne gefundenen Eigenschaften auf den Integralausdruck auf der rechten Seite von (3) anwendet und beachtet, daß $\mathfrak{G}$ eine Nullmenge ist.

Aus dem universellen Charakter der Klasseneinteilung folgt insbesondere, daß die von F. Vasilesco[3]) für die speziellen Potentialfunktionen hergeleiteten Sätze über die Verteilung der regulären und singulären Randpunkte auch für die allgemeine Differentialgleichung gelten.

(Eingegangen am 30. Oktober 1930.)

Über die Jacobische Kurve eines Systems von drei Kurven und den Begriff „Abhängigkeit“ bei homogenen Polynomen.

Von

H. Kapferer in Freiburg i. Br.

Zu je drei algebraischen ebenen Kurven $S_k = 0$ $(k = 1, 2, 3)$ gehört eine durch sie eindeutig bestimmte Jacobische Kurve $J = 0$. Sind die drei Kurven als homogene Polynome in drei Variablen x, y, z gegeben, mit den Ordnungszahlen m_1, m_2, m_3, so ist J die Funktionaldeterminante dieser drei Polynome. Wenn also J nicht gerade identisch verschwindet, so ist J ein homogenes Polynom von der Ordnung $m_1 + m_2 + m_3 - 3$. Die über die Kurve $J = 0$ bekannten Sätze sind zahlreich, aber nicht besonders tiefliegend. Sie betreffen großenteils den Fall, wo die drei Ordnungszahlen einander gleich sind, wo es sich also um ein Netz von Kurven handelt, von denen je drei linear unabhängige ein und dieselbe Jacobische Kurve liefern. Andere Sätze betreffen die Vielfachheit eines gemeinsamen Punktes aller drei Kurven in $J = 0$ und den besonderen Fall, wo die Kurve $J = 0$ eine der drei Kurven „berühren“ kann.

Die Ergebnisse der vorliegenden Abhandlung beruhen vor allem auf der Herbeiziehung eines scharfen Begriffes „Schnittpunktmultiplizität“. Wenn drei Kurven vorgegeben und damit auch implizite die Schnittpunkte π von je zweien festgelegt sind, wie verhält sich dann J zu diesen Punkten π, zunächst, wann geht J durch einen solchen Punkt π und wann nicht, vor allem aber, mit welcher Multiplizität kommt ein solcher Punkt π vor in jedem der drei Kurvenpaare $J = 0$, $S_k = 0$, $k = 1, 2, 3$? Für den Fall, daß π *einfacher* Punkt der irreduziblen Grundkurve $S_k = 0$ ist, werden diese Fragen vollständig gelöst.

Die Frage nach der Schnittpunktmultiplizität schließt aber noch eine andere sehr wesentliche in sich. Man spricht doch nur von Schnittpunktmultiplizität bei zwei Polynomen, wenn diese unter sich teilerfremd sind. Unter welchen Bedingungen ist J überhaupt teilerfremd zu S_k und wann

nicht? Diese Fragestellung ist schärfer als die Frage, wann verschwindet J identisch; denn wenn J identisch verschwindet, so ist ja a fortiori J auch teilbar durch S_k. Die Antwort, die im folgenden gegeben wird, ist derart, daß dabei der bekannte Satz über Abhängigkeit von homogenen Polynomen unter sich eine interessante Ergänzung erfährt.

Eine Zusammenfassung der Ergebnisse findet man in den „Hauptsätzen" I, II und III und in den aus ihnen folgenden Sätzen XI bis XIV, welche die Abhängigkeit von Polynomen betreffen.

Es ist vielleicht noch bemerkenswert, daß der Anlaß zur vorliegenden Abhandlung im Grunde ein zahlentheoretisches Problem war. Hierbei wurde ich ganz ungesucht auf Funktionaldeterminanten von drei ternären Formen geführt und auf die beiden genannten Fragen der Multiplizität und der Teilbarkeit von J durch eines der drei Polynome. Die nähere Erklärung hierzu geben die ersten Seiten meiner demnächst erscheinenden Abhandlung „Ein Beweis der Unmöglichkeit einer nicht linearen Korrespondenz zwischen doppelpunktfreien Kurven gleicher Ordnung $n > 3$". (Sitzungsberichte der Bayerischen Akademie 1931.)

Bei den Beweisen werden einige Tatsachen und Sätze aus früheren Arbeiten desselben Verfassers benützt. Diese sollen zunächst hier zitiert werden, und zwar gleich in der Form, wie sie im folgenden gebraucht werden, also nicht in der meist allgemeineren Fassung, in der sie in den zitierten Abhandlungen bewiesen worden sind. Wir fassen die notwendigen Tatsachen zusammen in die Sätze **K I**, **K II**, **K III**, **K IV**.

K I. (*Vier Eigenschaften der Schnittpunktmultiplizitätszahlen*). Es seien α, β, γ irgend drei Konstante, die nicht alle gleich Null sind; dann ist durch die Proportion $x:y:z = \alpha:\beta:\gamma$ ein Punkt π definiert. Jeder beliebige Punkt π hat in bezug auf jedes beliebige Paar teilerfremder homogener Polynome in x, y, z, etwa A und B, eindeutig eine Multiplizitätszahl, eine ganze Zahl, gleich oder größer als Null. Wir werden diese Zahl kurz mit (A, B) bezeichnen. Der Gebrauch dieses Symbols soll nur erlaubt sein, wenn die durch das Komma getrennten Polynome teilerfremd sind.

Die Eigenschaften, die wir nötig haben, sind folgende vier[1]): (Wir benötigen nur die *Tatsache* dieser Eigenschaften, nicht die tieferliegende Eigentümlichkeit, daß dieselben in ihrer Gesamtheit eine eindeutige Definition der Multiplizität ausmachen.)

---

[1]) Vgl. H. Kapferer, Axiomatische Begründung des Bézoutschen Satzes, Sitzungsberichte der Heidelberger Akademie 1927. Die Eigenschaften 1 und 4 sind trivial. Die Eigenschaft 2 beruht auf dem Produktsatz für Resultanten. (Vgl. v. d. Waerden, Moderne Algebra, II. Bd.) Die Eigenschaft 3 beruht ebenfalls auf einer allgemeinen Resultanteneigenschaft (vgl. F. Mertens, Zur Theorie der Elimination, § 8, Sitzungsberichte der Wiener Akademie 108 (1899)).

1. $(A, B) = 0$, wenn π *nicht* gemeinsame Nullstelle der beiden Polynome A und B ist.

2. $(A, C) + (B, C) = (AB, C)$.

3. $(A. B) = (A + tB, B)$; hierbei ist angenommen, daß die Ordnungszahl von A gleich oder größer als die von B, und daß diejenige von t gleich der Differenz der Ordnungszahlen von A und B ist.

4. $(A, B) = 1$ in dem besonderen Fall, daß A *und* B *lineare* Polynome sind, die *beide* den Punkt π als Nullstelle haben.

K II. (*Eine charakteristische Gleichung für μ-fache Schnittpunktmultiplizität*[2].) U und S seien zwei teilerfremde homogene Polynome in x, y, z von beliebiger Ordnung. π sei *einfache* Nullstelle von S, gegeben[3] durch $x:y:z = \alpha:\beta:1$. Bezüglich π sei $(U, S) \geqq k$; $k = 0$ ist zulässig; dann *existieren* zwei homogene Polynome A_k und C_k, welche die Kongruenz mod $[S]$ erfüllen:

$$UA_k \equiv (x - \alpha z)^k C_k,$$

und zwar kann man setzen

$$A_k = \{S(\alpha z, y, z) : (y - \beta z)\}^k,$$

so daß also $A_k = A_1^k$ ist. Dabei ist wesentlich, daß, wie aus der Definition von A_k hervorgeht, $A_k \not\equiv 0$ ist im Punkte π, sobald man voraussetzt[3], daß

$$S(\alpha z, y, z) \not\equiv o(y - \beta z)^2.$$

Außerdem gilt, gemäß **K I**, die Multiplizitätsrelation:

$$(U, S) = k + (C_k, S).$$

Falls also insbesondere $(U, S) = k$, so ist $(C_k, S) = 0$, d. h. C_k verschwindet *nicht* im Punkte π.

K III. Aus der Tatsache **K II** habe ich 1923 eine Regel zur praktischen Multiplizitätsbestimmung abgeleitet[4].

---

[2]) Vgl. hierzu H. Kapferer, Über die Multiplizität der Schnittpunkte von zwei algebraischen Kurven, Jahresber. d. Deutsch. Math.-Ver. **32** (1923/24). Am a. O. ist die Darstellung *nicht* homogen; doch läßt sich alles ohne Mühe in die homogene Form übertragen.

[3]) Würde der Punkt π in der Form $x:y:z = 1:\beta:\gamma$ oder in der Form $x:y:z = \alpha:1:\gamma$ gegeben sein, so würde an die Stelle von $(x - \alpha z)$ zu setzen sein $(y - \beta x)$ bzw. $(z - \gamma y)$; ferner an die Stelle von $S(\alpha z, y, z) \not\equiv o(y - \beta z)^2$ wäre zu setzen $S(x, y, \gamma x) \not\equiv o(y - \beta x)^2$ bzw. $S(\alpha y, y, z) \not\equiv o(z - \gamma x)^2$. Sobald also π *einfacher* Punkt von S ist, so ist es nur eine Sache der geeigneten Bezeichnung, zu erreichen, daß A_k *nicht* im Punkte π verschwindet.

[4]) Vgl. [2]) sowie § 2 der Abhandlung H. Kapferer, Schnittpunktsysteme mit vorgeschriebenen Multiplizitätszahlen, Sitzungsber. der Heidelberger Akademie 1930.

π sei gemeinsame Nullstelle von A und S, also $(A, S) = \mu > 0$; insbesondere in S sei π *einfache* Nullstelle; S selbst sei irreduzibel.

π sei gegeben in der Form $x : y : z = \alpha : \beta : 1$. Dann gilt der Satz:

$$\left(\frac{\partial(A,S)}{\partial(x,y)}, S\right) = \mu - 1,$$

wobei $\frac{\partial(A,S)}{\partial(x,y)} = \frac{\partial A}{\partial x}\frac{\partial S}{\partial y} - \frac{\partial A}{\partial y}\frac{\partial S}{\partial x}$ bedeutet[5]).

K IV. (Aus dem Ideenkreis des Noetherschen Fundamentalsatzes.) Es handelt sich um den folgenden, in gewissem Sinne einfachsten Fall der a. a. O.[6]) behandelten Problemstellung.

Zwei teilerfremde ternäre Formen A und B in x, y, z seien fest gegeben. Jede gemeinsame Nullstelle von A und B sei *in B einfache* Nullstelle. Die Bedingung, welche einem dritten Polynom C durch die Forderung auferlegt wird, daß C in die Form gebracht werden soll

$$C = \lambda A + \mu B,$$

wo λ und μ ebenfalls ternäre Formen sind, ist *äquivalent* so vielen Schnittpunktmultiplitätsbedingungen, als es *verschiedene* Schnittpunkte von $A = 0$, $B = 0$ gibt. Wenn nämlich π irgendeine gemeinsame Nullstelle von A und B bedeutet, so muß — C als teilerfremd zu B vorausgesetzt — die Multiplizität des Punktes π im Kurvenpaar $C = 0$, $B = 0$ *mindestens so groß* sein, wie im Kurvenpaar $A = 0$, $B = 0$.

Das Wesentliche des Satzes besteht darin, daß die angegebenen Bedingungen hinreichend sind für die gewünschte Darstellung von C.

Unter Satz I fassen wir vier einfache Tatsachen zusammen, deren Inhalt fast unmittelbar aus der Definition der Funktionaldeterminante von drei ternären Formen folgt und die daher keines besonderen Beweises bedürfen. Als Symbol für die Funktionaldeterminante von f_1, f_2, f_3 in x, y, z benützen wir, wie üblich, $\frac{\partial(f_1, f_2, f_3)}{\partial(x, y, z)}$.

---

[5]) Würde der Punkt π in der Form $x : y : z = 1 : \beta : \gamma$ oder in der Form $x : y : z = \alpha : 1 : \gamma$ gegeben sein, so würde an die Stelle von $\frac{\partial(A,S)}{\partial(x,y)}$ zu setzen sein $\frac{\partial(A,S)}{\partial(y,z)}$ bzw. $\frac{\partial(A,S)}{\partial(z,x)}$.

Davon, daß der Satz **K III** *umkehrbar* ist — daß also μ sukzessive numerisch bestimmt werden kann —, machen wir im folgenden keinen Gebrauch.

[6]) Vgl. H. Kapferer, Notwendige und hinreichende Multiplizitätsbedingungen zum Noetherschen Fundamentalsatz der algebraischen Funktionen, Sitzungsberichte der Heidelberger Akademie 1927, sowie Math. Annalen 97 (1927).

a) $$\frac{\partial(A+B,C,S)}{\partial(x,y,z)} = \frac{\partial(A,C,S)}{\partial(x,y,z)} + \frac{\partial(B,C,S)}{\partial(x,y,z)}.$$

b) $$\frac{\partial(AB,C,S)}{\partial(x,y,z)} = A\frac{\partial(B,C,S)}{\partial(x,y,z)} + B\frac{\partial(A,C,S)}{\partial(x,y,z)}.$$

c) Aus Satz I b folgt insbesondere

$$\frac{\partial(AS,C,S)}{\partial(x,y,z)} \equiv o\,[S].$$

d) Falls $A \equiv B \bmod [S]$, so folgt aus Satz I a und I c die Kongruenz mod $[S]$:

$$\frac{\partial(A,C,S)}{\partial(x,y,z)} \equiv \frac{\partial(B,C,S)}{\partial(x,y,z)}.$$

Satz II. *Es seien f und g zwei ternäre Formen gleicher Ordnung m, und t eine ternäre Form der Ordnung d. Ferner sei*

$$F = tf \quad \text{und} \quad G = tg.$$

S sei eine Form der Ordnung s. Dann gilt die Kongruenz mod $[S]$:

$$m\frac{\partial(F,G,S)}{\partial(x,y,z)} \equiv (m+d)\,t^2\frac{\partial(f,g,S)}{\partial(x,y,z)}.$$

Denn

$$J = \frac{\partial(F,G,S)}{\partial(x,y,z)} = \begin{vmatrix} tf_x + ft_x, & tf_y + ft_y, & tf_z + ft_z \\ tg_x + gt_x, & tg_y + gt_y, & tg_z + gt_z \\ S_x, & S_y, & S_z \end{vmatrix}$$

$$= t^1\left\{g\frac{\partial(f,t,S)}{\partial(x,y,z)} + f\frac{\partial(t,g,S)}{\partial(x,y,z)}\right\} + t^2\frac{\partial(f,g,S)}{\partial(x,y,z)}.$$

Dies läßt sich noch umformen auf Grund der Relationen:

$$\begin{aligned} mf &= xf_x + yf_y + zf_z, \\ my &= xg_x + yg_y + zg_z, \\ dt &= xt_x + yt_y + zt_z, \\ sS &= xS_x + yS_y + zS_z, \end{aligned}$$

denn hiernach verschwindet nachstehende Determinante vierten Grades identisch:

$$\begin{vmatrix} mf, & f_x, & f_y, & f_z \\ mg, & g_x, & g_y, & g_z \\ dt, & t_x, & t_y, & t_z \\ sS, & S_x, & S_y, & S_z \end{vmatrix}.$$

Daraus folgt aber durch Entwicklung nach der ersten Kolonne die Kongruenz mod $[S]$:

$$m\left\{g\frac{\partial(f,t,S)}{\partial(x,y,z)} + f\frac{\partial(t,g,S)}{\partial(x,y,z)}\right\} \equiv dt^1\frac{\partial(f,g,S)}{\partial(x,y,z)}.$$

Hieraus und aus dem obigen Ausdruck für J folgt sofort die in Satz II angegebene Kongruenz.

Satz III. *Seien wieder f und g zwei ternäre Formen gleicher Ordnung m; S von beliebiger Ordnung s. A_k sei eine solche der Ordnung d, φ_k und ψ_k zwei solche von gleicher Ordnung q. Zwischen den genannten fünf Polynomen mögen folgende Kongruenzen* mod $[S]$ *bestehen:*

a) $$fA_k \equiv (x-\alpha z)^k \varphi_k, \quad gA_k \equiv (x-\alpha z)^k \psi_k.$$

Dann, behaupten wir, erfüllt das Polynom

$$B_{2k} = A_k^2 \frac{q(m+d)}{m(q+k)}$$

nachstehende Kongruenz mod $[S]$:

b) $$B_{2k} \frac{\partial(f,g,S)}{\partial(x,y,z)} \equiv (x-\alpha z)^{2k} \frac{\partial(\varphi_k, \psi_k, S)}{\partial(x,y,z)}.$$

Beweis. Zunächst gilt wegen Satz I d jedenfalls mod $[S]$:

$$\frac{\partial(fA_k, gA_k, S)}{\partial(x,y,z)} \equiv \frac{\partial((x-\alpha z)^k \varphi_k, (x-\alpha z)^k \psi_k, S)}{\partial(x,y,z)}.$$

Wir wenden zweimal den Satz II an, nämlich auf jede der beiden letztgenannten Funktionaldeterminanten, und erhalten so mod $[S]$:

$$m \frac{\partial(fA_k, gA_k, S)}{\partial(x,y,z)} \equiv (m+d) A_k^2 \frac{\partial(f,g,S)}{\partial(x,y,z)},$$

$$q \frac{\partial((x-\alpha z)^k \varphi_k, (x-\alpha z)^k \psi_k, S)}{\partial(x,y,z)} \equiv (q+k)(x-\alpha z)^{2k} \frac{\partial(\varphi_k, \psi_k, S)}{\partial(x,y,z)}.$$

Eine einfache Kombination der drei letztgenannten Kongruenzen führt schließlich zu der Kongruenz des Satzes III, die zu beweisen war.

Auf Satz III baut sich alles Weitere auf, bei gleichzeitiger Verschärfung der Voraussetzungen.

Voraussetzungen des Satzes IV: Es seien f und g Formen gleicher Ordnung m; S irreduzibel; $f \not\equiv o[S]$, $g \not\equiv o[S]$. Ein Punkt π von der Form $x:y:z = \alpha:\beta:1$ sei *einfacher* Punkt von S, so daß also *wenigstens eine* der Ungleichungen

$$S(x, \beta z, z) \not\equiv o(x-\alpha z)^2 \quad \text{und} \quad S(\alpha z, y, z) \not\equiv o(y-\beta z)^2$$

erfüllt ist. Wir nehmen an — mit Rücksicht auf den Wortlaut von K II —, daß die *zweite* der genannten Ungleichungen erfüllt ist. Es sei bezüglich π

$$(f,S) \geqq k \geqq 0, \quad (g,S) \geqq k \geqq 0.$$

Unter diesen Voraussetzungen gilt der

Satz IV. *Die Polynome A_k, φ_k, ψ_k, B_{2k} des Satzes* III *existieren, und zwar so, daß A_k, und somit auch B_{2k}, im Punkte π nicht verschwinden.*

Der Beweis ergibt sich unmittelbar aus Satz III, in Verbindung mit **K II**. Daraus ziehen wir für das Weitere grundlegende Folgerungen. Wir setzen zur Abkürzung

$$\frac{\partial(f,g,S)}{\partial(x,y,z)} = J; \qquad \frac{\partial(\varphi_k,\psi_k,S)}{\partial(x,y,z)} = J_k$$

und behaupten:

Satz IVa. *Unter den Voraussetzungen des Satzes* IV *gilt*:

$$J \equiv o[S] \text{ dann und nur dann, wenn } J_k \equiv o[S].$$

Denn S ist irreduzibel, daher $B_{2k} \not\equiv o[S]$, da ja B_{2k} in einer Nullstelle von S *nicht* verschwindet, nämlich in π. Ferner ist $x - \alpha z \neq S$, weil sonst — entgegen der Voraussetzung — $S(\alpha z, y, z) \equiv o(y - \beta z)^2$ wäre.

Wir behaupten ferner:

Satz IVb. *Unter den Voraussetzungen des Satzes* IV *und der weiteren Voraussetzung* $J \not\equiv o[S]$ *gilt bezüglich Punkt* π:

$$(J, S) = 2k + (J_k, S).$$

Denn aus Satz IIIb, in Verbindung mit **K I**, folgt, falls $J \not\equiv o[S]$:

$$(J, S) = ((x - \alpha z)^{2k}, S) + (J_k, S).$$

Es ist aber $(x - \alpha z, S) = 1$ bezüglich π infolge der Voraussetzung $S(\alpha z, y, z) \not\equiv o(y - \beta z)^2$.

Zusatz. Aus den Anmerkungen zu **K II** und **K III** geht hervor, daß die Sätze IV, IVa, IVb in ganz analoger Form auszusprechen sind, wenn der *einfache* Punkt π von S *nicht* in der in Satz III genannten Form, sondern in einer der beiden anderen noch möglichen Formen gegeben ist. Die Formulierung der betreffenden Sätze und ihr Beweis erfordert nur geeignete Vertauschungen der Größen α, β, γ unter sich, und der Variablen x, y, z unter sich.

Satz V. *A, B, S seien drei ternäre Formen; S irreduzibel; $A \not\equiv o[S]$; $B \not\equiv o[S]$; A und B gleicher Ordnung. Ein Punkt π sei einfacher Punkt von S. Bezüglich π sei*

$$(A, S) = \mu; \qquad (B, S) = \nu; \qquad \mu \neq \nu.$$

Dann behaupten wir

a) $$A + B \not\equiv o[S],$$

b) $$(A + B, S) = \textit{Minimum } \mu, \nu.$$

Beweis. Wir nehmen zunächst an, π sei von der Form $x : y : z = \alpha : \beta : 1$ und ferner, daß $S \neq (x - \alpha z)$ ist. Sei etwa $\mu > \nu$, dann existieren nach **K II** vier im Punkte π nicht verschwindende Polynome $C_\mu, C_\nu, \varphi_\mu, \psi_\nu$

welche nachstehende Kongruenzen mod $[S]$ erfüllen:

$$A\,C_\mu \equiv (x-\alpha z)^\mu \varphi_\mu,$$
$$B\,C_\nu \equiv (x-\alpha z)^\nu \psi_\nu.$$

Hierbei ist nach **K II** $C_\mu = C_1^\mu$; $C_\nu = C_1^\nu$; daher folgt

$$(A+B)\,C_1^\mu \equiv (x-\alpha z)^\nu \cdot \chi \mod [S],$$

wobei

$$\chi = (x-\alpha z)^{\mu-\nu}\varphi_\mu + C_1^{\mu-\nu}\cdot\psi_\nu$$

ist. Da nun C_1 und ψ_ν beide *nicht* in π verschwinden, der *erste* Summand von χ aber in π verschwindet, wegen $\mu-\nu>0$, so verschwindet χ *nicht* im Punkte π, also ist a fortiori $\chi \not\equiv o[S]$, also auch $A+B \not\equiv o[S]$, weil $x-\alpha z \not\equiv S$ nach Voraussetzung. Daher folgt nach **K I**

$$(A+B, S) = ((x-\alpha z)^\nu, S) = \nu.$$

Aus den Anmerkungen zu **K II** und **K III** sowie aus dem Zusatz zu Satz IV geht hervor, daß der Beweis *analog* verläuft, wenn π nicht in der angegebenen Form, sondern in einer der beiden anderen noch möglichen Formen gegeben ist. Dann ist aber Satz V allgemein bewiesen.

Satz VI. *Sei S irreduzibel; $f \not\equiv o[S]$; $g \not\equiv o[S]$; f und g gleicher Ordnung m; π sei gemeinsame Nullstelle von f und S, aber nicht von g; π sei einfache Nullstelle von S; also $(f,S)>0$; $(g,S)=0$; dann behaupten wir:*

a) $$J \not\equiv o[S],$$

b) $$(J,S) = (f,S)-1.$$

Dieser Satz VI ist als Spezialfall im Satz VIII enthalten, muß aber vorausbewiesen werden, weil er beim Beweis von Satz VIII gebraucht wird. Der Beweis beruht wesentlich auf der Regel des Satzes **K III**.

Aus der Identität

$$z\cdot J = \begin{vmatrix} f_x, & f_y, & m f \\ g_x, & g_y, & m g \\ S_x, & S_y, & s S \end{vmatrix}$$

folgt die Kongruenz mod $[S]$:

$$z\cdot J \equiv m\left\{ f\frac{\partial(g,S)}{\partial(x,y)} + g\frac{\partial(S,f)}{\partial(x,y)}\right\}.$$

Nun ist aber nach der Regel **K III**

$$\left(\frac{\partial(f,S)}{\partial(x,y)}, S\right) = (f,S)-1;$$

also, wegen $(g,S)=0$, auch

$$\left(g\frac{\partial(f,S)}{\partial(x,y)}, S\right) = (f,S)-1.$$

Andererseits ist

$$\left(f\frac{\partial(g,S)}{\partial(x,y)},S\right)\geqq(f,S);$$

daher folgt aus Satz V

$$(zJ,S)=(f,S)-1$$

und, wegen $(z,S)=0$, falls π von der Form $x:y:z=\alpha:\beta:1$ ist, folgt

$$(J,S)=(f,S)-1.$$

Aus den Anmerkungen zu **K II** und **K III** geht wieder hervor, daß der Beweis analog verläuft, wenn π in einer der beiden anderen noch möglichen Formen gegeben ist. Man hätte, statt von zJ, auszugehen von xJ bzw. von yJ.

Satz VII (= Hilfssatz über J_k). *Es sei* $(f,S)>k$ *und* $(g,S)=k$, *und es seien im übrigen die Voraussetzungen des Satzes IV erfüllt. Dann behaupten wir*:

a) $J_k\not\equiv o[S]$,

b) $(J_k,S)=(f,S)-(g,S)-1.$

Denn nach **K I** und Satz III a ist

$$(f,S)=k+(\varphi_k,S),$$
$$(g,S)=k+(\psi_k,S).$$

Infolge der Voraussetzungen $(f,S)>k$ und $(g,S)=k$ wird also $(\varphi_k,S)>0$ und $(\psi_k,S)=0$. Demnach kann man auf das Tripel φ_k,ψ_k,S den Satz VI anwenden. Es folgt auf diese Weise:

1. $J_k\not\equiv o[S]$,

2. $(J_k,S)=(\varphi_k,S)-1,$

was zu beweisen war; denn es ist ja $(f,S)=k+(\varphi_k,S)=(g,S)+(\varphi_k,S)$.

Satz VI ist ein Spezialfall des Satzes VIII, den wir nun beweisen wollen.

Satz VIII[7]). *Sei* S *irreduzibel*; $f\not\equiv o[S]$; $g\not\equiv o[S]$; f *und* g *gleicher Ordnung*; π *sei einfache Nullstelle von* S, *für welche gelten soll*

$$(f,S)\neq(g,S).$$

Dann behaupten wir

a) $J\not\equiv o[S]$,

b) $(J,S)=(f,S)+(g,S)-1.$

---

[7]) Die Vereinigung der Sätze VIII und X liefert sofort folgenden übersichtlichen Zusatz (auf den sich die auf Seite 555 angekündigte Abhandlung stützen wird):

f und g seien gleicher Ordnung; S irreduzibel; $J\not\equiv o[S]$. Dann gilt für *jede einfache* Nullstelle von S die Ungleichung

$$(J,S)\geqq 2\mu,$$

wo μ die kleinere der beiden Zahlen (f,S) und (g,S) ist.

Beweis von Satz VIIIa. Wegen $(f, S) \neq (g, S)$ ist $J_k \not\equiv o[S]$ nach Satz VIIa. Aus $J_k \not\equiv o[S]$ folgt aber $J \not\equiv o[S]$ nach Satz IVa.

Beweis von Satz VIIIb. Nach Satz IVb ist

$$(J, S) = 2k + (J_k, S).$$

Nach Voraussetzung ist $(f, S) \neq (g, S)$, also etwa $(f, S) > (g, S) = k$. Dann ist nach Satz VIIb

$$(J_k, S) = (f, S) - (g, S) - 1$$

und also

$$(J, S) = 2k + (J_k, S) = (f, S) + (g, S) - 1.$$

Satz IX. *Sei S irreduzibel; $f \not\equiv o[S]$; $g \not\equiv o[S]$; f und g gleicher Ordnung; jeder Schnittpunkt von $f = 0$, $S = 0$ und jeder Schnittpunkt von $g = 0$, $S = 0$ sei einfacher Punkt von S. Dann gilt $J \equiv o[S]$ dann und nur dann, wenn es eine Konstante c gibt, so daß $f \equiv c \cdot g \bmod [S]$ ist.*

Beweis. Zunächst ist leicht zu zeigen, daß aus $f \equiv cg \bmod S$ folgt $J \equiv o[S]$. Das ist ja eine bloße Folge aus Satz Id.

Das Wesentliche ist die Umkehrung; die folgt so:

Wegen Satz VIIIa kann höchstens dann $J \equiv o[S]$ austreten, wenn für *jeden einzelnen* gemeinsamen Punkt der *drei* Polynome f, g, S gilt

$$(f, S) = (g, S).$$

In diesem Falle aber folgt aus KIV, daß es eine Konstante c gibt, so daß $f \equiv cg \bmod [S]$ ist.

Satz X[7]). *Sei $J \not\equiv o[S]$; π sei gemeinsame Nullstelle der drei Polynome f, g, S; π sei einfache Nullstelle von S, und es sei insbesondere*

$$(f, S) = (g, S),$$

dann gilt

$$(J, S) \geqq (f, S) + (g, S).$$

Beweis folgt unmittelber aus Satz IVb.

Bisher war stets angenommen worden, daß f und g von gleicher Ordnung sind. Wir wollen jetzt diese Beschränkung beseitigen und kommen auf diese Weise zu den drei Hauptsätzen dieser Abhandlung.

Sei f von der Ordnung m_1, g von der Ordnung m_2; dann setzen wir

$$f^{m_2} = \bar{f}; \quad g^{m_1} = \bar{g},$$

so daß $\bar{f}$ und $\bar{g}$ von *gleicher* Ordnung $m_1 m_2$ sind. Wir setzen ferner zur Abkürzung

$$\frac{\partial(\bar{f}, \bar{g}, S)}{\partial(x, y, z)} = \bar{J}.$$

Dann folgt, durch wiederholte Anwendung des Satzes Ib,

$$\bar{J} = m_1 m_2 f^{m_2 - 1} g^{m_1 - 1} J.$$

Es ist also $\bar{J} \equiv o[S]$ nur dann, wenn entweder $f \equiv o[S]$ oder $g \equiv o[S]$ oder $J \equiv o[S]$.

Nun wenden wir Satz IX an auf das Tripel $\bar{f}, \bar{g}, S$ und erhalten, wenn wir noch beachten, daß die Kongruenz $f^{m_2} \equiv c g^{m_1} \bmod [S]$ bei irreduziblem S ersetzt werden kann durch die andere

$$f^{\frac{m_2}{d}} \equiv c_3 g^{\frac{m_1}{d}} \bmod [S],$$

wo $d = (m_1, m_2)$ der größte gemeinsame Teiler von m_1 und m_2 ist, den folgenden

Hauptsatz I. *Sei S irreduzibel; $f \not\equiv o[S]$; $g \not\equiv o[S]$; alle drei beliebiger Ordnung; f von der Ordnung m_1, g von der Ordnung m_2. Jeder Schnittpunkt von $f = 0$, $S = 0$ und jeder Schnittpunkt von $g = 0$, $S = 0$ sei einfacher Punkt von $S = 0$. Dann gilt $J \equiv o[S]$ dann und nur dann, wenn eine Konstante c existiert, so daß*

$$f^{\frac{m_2}{(m_1, m_2)}} \equiv c g^{\frac{m_1}{(m_1, m_2)}} \bmod [S].$$

Wir kommen zum Hauptsatz II, indem wir Satz VIII anwenden auf das Tripel $\bar{f}, \bar{g}, S$.

Wenn nämlich $J \not\equiv o[S]$, $f \not\equiv o[S]$, $g \not\equiv o[S]$, so folgt nach K I

$$\begin{aligned}(\bar{J}, S) &= (f^{m_2 - 1}, S) + (g^{m_1 - 1}, S) + (J, S) \\ &= (m_2 - 1)(f, S) + (m_1 - 1)(g, S) + (J, S);\end{aligned}$$

$$(\bar{f}, S) = m_2 (f, S), \quad (\bar{g}, S) = m_1 (g, S).$$

Daher bedeutet die Relation des Satzes VIII:

$$(\bar{J}, S) \equiv (\bar{f}, S) + (\bar{g}, S) - 1$$

mit der Bedingung $(\bar{f}, S) \neq (\bar{g}, S)$ soviel wie

$$(J, S) = (f, S) + (g, S) - 1$$

mit der Bedingung $m_2 (f, S) \neq m_1 (g, S)$.

Wir haben also bewiesen:

Hauptsatz II. *Sei S irreduzibel; $f \not\equiv o[S]$; $g \not\equiv o[S]$; π sei einfache Nullstelle von S. (f, S) und (g, S) seien nicht beide Null bezüglich π und erfüllen die Ungleichung*

$$(f, S) : (g, S) \neq m_1 : m_2.$$

Dann gilt

a) $J \not\equiv o[S]$,

b) $(J, S) = (f, S) + (g, S) - 1$.

Auf ganz analoge Weise findet man durch Anwendung des Satzes X auf das Tripel $\bar{f}, \bar{g}, S$, indem man die angegebenen Beziehungen zwischen $(\bar{f}, S)$, $(\bar{g}, S)$, $(\bar{J}, S)$ einerseits und (f, S), (g, S), (J, S) andererseits benützt und beachtet, daß die Forderung

$$(\bar{f}, S) = (\bar{g}, S)$$

soviel bedeutet wie $m_2(f, S) \equiv m_1(g, S)$ den folgenden

Hauptsatz III. *Sei S irreduzibel; $J \not\equiv o[S]$; π sei einfache Nullstelle von S; (f, S) und (g, S), bezüglich π genommen, erfüllen die Gleichung:*

$$(f, S) : (g, S) = m_1 : m_2.$$

Dann gilt

$$(J, S) \geqq (f, S) + (g, S).$$

Beispiele zu den Hauptsätzen II und III.

Zu Hauptsatz II. a) $f = x^5$; $g = y^7$; $S = xz + y^2$. Bezüglich Punkt $x : y : z = 0 : 0 : 1$ gilt

$$(f, S) = 10; \quad (g, S) = 7; \quad (f, S) : (g, S) + m_1 : m_2 = 5 : 7;$$

$$J = \begin{vmatrix} 5x^4, & 0, & 0 \\ 0, & 7y^6, & 0 \\ z, & 2y, & x \end{vmatrix} = 35 x^5 y^6,$$

also $(J, S) = (x^5, S) + (y^6, S) = 10 + 6 = 16$.

In der Tat ist $16 = (f, S) + (g, S) - 1$.

b) $f = x^n$; $g = y^{2n}$; $S = xz + y^2$; $(f, S) = 2n$; $(g, S) = 2n$; $J = 2n^2 x^n y^{2n-1}$; $(J, S) = (x^n, S) + (y^{2n-1}, S) = 2n + 2n - 1 = 4n - 1$.

In der Tat ist $4n - 1 = (f, S) + (g, S) - 1$.

Zu Hauptsatz III. a) $f = x^5$; $g = y^5$; $S = (x + y)z + x^2 + y^2$. Bezüglich Punkt $x : y : z = 0 : 0 : 1$ gilt

$$(f, S) = 5; \quad (g, S) = 5, \quad \text{also} \quad (f, S) = (g, S); \quad m_1 = m_2;$$

$J = 25 x^4 y^4 (x + y)$; also $(J, S) = (x^4, S) + (y^4, S) + (x + y, S) = 4 + 4 + 2 = 10$.

In Übereinstimmung mit Hauptsatz III ist $10 = (f, S) + (g, S)$.

b) Nach Hauptsatz III kann auch $(J, S) > (f, S) + (g, S)$ werden. Auch hierfür ein Beispiel:

$$f = x^3; \quad g = y^2; \quad S = z^2(x + y) + x^3 + y^3;$$

$$(f, S) = 2; \quad (g, S) = 2; \quad J = 8xyz(x + y), \quad \text{also}$$

$(J, S) = (x, S) + (y, S) + (x + y, S) = 1 + 1 + 3 = 5 = 1 + (f, S) + (g, S)$.

Die Voraussetzungen der drei Hauptsätze haben ihre Ursache in den beim Beweis benützten Sätzen **K II, KIII, K IV.** Es ist vielleicht erwünscht,

durch direkte Beispiele zu zeigen, *wie die Hauptsätze bei Abschwächung der Voraussetzungen falsch werden können.*

Wenn man z. B. für S das Quadrat eines Polynoms, etwa $S = T^2$ setzt, so werden alle drei Hauptsätze falsch.

Denn nach Satz Ib wird $J = \frac{\partial(f, g, S)}{\partial(x, y, z)} = 2T\frac{\partial(f, g, T)}{\partial(x, y, z)}$. Es ist also J gar nicht teilerfremd zu S, und man kann daher gar nicht von Multiplizität (J, S) reden, die Hauptsätze II und III werden bedeutungslos. Aber auch Hauptsatz I wird falsch; denn man kann $J \equiv o[S]$ erreichen, ohne daß eine Konstante c existiert, so daß $f \equiv cg \mod [S]$ ist. Man wähle z. B. f und g so, daß $f \equiv cg \mod [T^1]$, aber $f \not\equiv cg \mod [T^2]$; dann folgt aus Satz Id, daß $\frac{\partial(f, g, T)}{\partial(x, y, z)} \equiv o[T]$; also wird $J = \frac{\partial(f, g, S)}{\partial(x, y, z)} \equiv o[T^2] \equiv o[S]$.

Wie schon ein Doppelpunkt von S den Hauptsatz II unrichtig machen kann, zeigt folgendes Beispiel:

$$f = x^2; \quad g = xy + z^2; \quad S = x^3 + y^2.$$

$$(f, S) = 4; \quad (g, S) = 0; \quad \text{also} \quad (f, S) \neq (g, S); \quad m_1 = m_2;$$

$$J = -8xyz; \quad (J, S) = (x, S) + (y, S) = 2 + 2 = 4,$$

während nach Hauptsatz II sein müßte $(J, S) = 4 + 0 - 1 = 3$.

Folgerungen bezüglich der „Abhängigkeit" von homogenen Polynomen unter sich.

Drei homogene Polynome f_1, f_2, f_3 in x_1, x_2, x_3 werden als voneinander abhängig bezeichnet, wenn es ein Polynom $G(y_1, y_2, y_3)$ gibt, das in y_1, y_2, y_3 nicht identisch verschwindet, während

$$G(f_1, f_2, f_3) = 0$$

eine Identität in x_1, x_2, x_3 ist.

Für eine solche Abhängigkeit ist bekanntlich das identische Verschwinden der Funktionaldeterminante J von f_1, f_2, f_3 charakteristisch. Der wesentliche und auch im Beweis schwierigere Teil dieses Satzes besteht darin, daß das identische Verschwinden von J hinreicht für eine solche Abhängigkeit[8]).

Nun ziehen wir den Hauptsatz I hinzu.

Satz XI. *f_3 sei irreduzibel. Die Schnittpunkte von $f_1 = 0$, $f_3 = 0$ und ebenso die Schnittpunkte von $f_2 = 0$, $f_3 = 0$ seien einfache Punkte von $f_3 = 0$; $f_1 \not\equiv o[f_3]$; $f_2 \not\equiv o[f_3]$. Dann folgt aus dem identischen Ver-*

[8]) Eine präzise Beweismethode — für k beliebige Polynome mit k Variabeln — findet man in Perrons Lehrbuch der Algebra, 1, § 27. Man vgl. auch desselben Verfassers „Bemerkungen zur Algebra", Sitzungsberichte der Münchener Akademie 1924. Dort sind auch verschiedene Verallgemeinerungen behandelt.

schwinden von J — weil alle Voraussetzungen des Hauptsatzes I *erfüllt sind —, daß es eine Konstante c gibt, so das*

$$f_1^{\frac{m_2}{d}} = c f_2^{\frac{m_1}{d}} \mod [f_3] \quad \textit{ist} \qquad [d = (m_1, m_2)].$$

Diese letztere Kongruenz ist, wie mir scheint, viel mehr sagend als die obige Aussage, daß ein im übrigen ganz unbekanntes Polynom $G(y_1, y_2, y_3)$ existiert mit der angegebenen Eigenschaft. Wenn man sich schließlich auf drei Formen *gleicher* Ordnung beschränkt, so bedeutet die Kongruenz des Satzes XI sogar eine *lineare* Abhängigkeit. Wir haben also unter andern folgende beide Sätze:

Satz XII. *Wenn f_1, f_2, f_3 drei Formen gleicher Ordnung sind, f_3 frei von vielfachen Punkten, also a fortiori irreduzibel; $f_1 \not\equiv o[f_3]$; $f_2 \not\equiv o[f_3]$. Dann sind die drei Formen entweder überhaupt unabhängig voneinander, oder es besteht lineare Abhängigkeit*

$$c_1 f_1 + c_2 f_2 + c_3 f_3 = 0.$$

Satz XIII. *Wenn zwischen drei homogenen Polynomen gleicher Ordnung eine nicht lineare Abhängigkeit besteht, und je zwei derselben teilerfremd sind, so muß jedes derselben je wenigstens einen vielfachen Punkt besitzen.*

Beispiel zu Satz XIII. Zwischen $f_1 = 2x_1x_3$; $f_2 = x_1^2 - x_2^2$; $f_3 = x_1^2 + x_2^2$ besteht die quadratische Abhängigkeit $f_1^2 + f_2^2 = f_3^2$, aber sicher keine lineare Abhängigkeit. In der Tat besitzen alle drei Polynome einen vielfachen Punkt, nämlich $x_1 : x_2 : x_3 = 0 : 0 : 1$ als Doppelpunkt.

In Satz XI ist unter andern auch der folgende

Satz XIV enthalten. *Die Voraussetzungen des Satzes* XI *seien erfüllt. Außerdem sei*

$$m_3 > \frac{m_1 \cdot m_2}{(m_1, m_2)}.$$

Dann folgt aus dem identischen Verschwinden von J nicht nur die Kongruenz des Satzes XI, *sondern sogar direkte Gleichheit*

$$f_1^{\frac{m_2}{d}} = c g^{\frac{m_1}{d}}.$$

(Eingegangen am 20. Januar 1931.)

Axiomatik der Wahrscheinlichkeitsrechnung.

Von

Hans Reichenbach in Berlin.

I. Grundlagen der mathematischen Wahrscheinlichkeitsrechnung.

§ 1.

Problemstellung.

In der axiomatischen Begründung der Wahrscheinlichkeitsrechnung sind bisher zwei verschiedene Wege beschritten worden. Einerseits hat man die klassische Wahrscheinlichkeit behandelt, in welcher die Wahrscheinlichkeit als eine Eigenschaft von Ereignissen erscheint; Axiomensysteme sind hier meist nur in provisorischer Form aufgestellt worden, da man nie untersucht hat, ob in den Schlüssen nicht noch mehr Voraussetzungen benutzt werden. Andererseits hat man die statistische Wahrscheinlichkeitsrechnung, in der die Wahrscheinlichkeit als Häufigkeit gedeutet wird, untersucht; in den Arbeiten von v. Mises [1]), Dörge [2]), Tornier [3]) sind auf diesem Gebiet bereits axiomatische Untersuchungen durchgeführt worden, in denen die notwendigen Anforderungen an axiomatische Strenge berücksichtigt worden sind.

Die folgende Axiomatik will beide Richtungen in sich vereinen. Sie geht von der klassischen oder formalen Wahrscheinlichkeitsrechnung aus,

---

[1]) R. v. Mises, Grundlagen der Wahrscheinlichkeitsrechnung, Math. Zeitschr. **5** (1919), S. 52.

[2]) Karl Dörge, Zu der von R. v. Mises gegebenen Begründung der Wahrscheinlichkeitsrechnung, Math. Zeitschr. **32** (1930), S. 232.

[3]) Erhard Tornier, Eine neue Grundlegung der Wahrscheinlichkeitsrechnung, Zeitschr. f. Physik **63** (1930), S. 697; Die Axiome der Wahrscheinlichkeitsrechnung, Crelles Journal 1930.

n der von der Häufigkeitsdeutung kein Gebrauch gemacht wird, und stellt ür diese ein Axiomensystem nach strengen axiomatischen Methoden auf. Dabei wird ein Kalkül entwickelt werden, der ein besonderes Zeichen für lie Wahrscheinlichkeit erhält; dieser Kalkül ist dem üblichen Logikkalkül ngepaßt und wird in Verbindung mit diesem benutzt. Die Wahrschein-ichkeitsrechnung wird damit dem Logikkalkül eingeordnet; sie gewinnt ine ähnliche Form wie eine in den Zeichen des Logikkalküls einschließ-ich besonderer Grundzeichen geschriebene Geometrie. Sodann aber wird ezeigt werden, daß das aufgestellte Axiomensystem zugleich als Axiomen-ystem der statistischen Wahrscheinlichkeitsrechnung aufgefaßt werden kann. n dieser Tatsache sehen wir eines der wichtigsten Ergebnisse der vor-iegenden Untersuchung; zugleich stellt sich dabei heraus, daß auch die tatistische Wahrscheinlichkeitsrechnung durch ein einfaches und durch-ichtiges Axiomensystem erfaßt werden kann, ohne daß paradoxe Begriffs-ildungen auftreten.

Das Problem der Mathematisierung der Wahrscheinlichkeitsrechnung st damit gelöst; der Wahrscheinlichkeitsbegriff erscheint als ein strenger nathematischer Begriff, und was er sonst noch an Schwierigkeiten enthält, eht jedenfalls in die *mathematische* Wahrscheinlichkeitsrechnung nicht in. Diese Schwierigkeiten, die man als Unbestimmtheit und Unsicherheit ler Wahrscheinlichkeitsaussage empfindet, treten vielmehr erst mit der Anwendung des Wahrscheinlichkeitsbegriffs auf Naturdinge ein, also erst ür die angewandte Wahrscheinlichkeitsrechnung. Wir werden deshalb in inem III. Teil auf die Frage der Anwendung eingehen. Die genannten Schwierigkeiten klären sich dabei auf eine sehr eigenartige Weise auf. Es wird sich nämlich zeigen, daß jede Anwendung des Wahrscheinlichkeits-egriffs auf wirkliche Dinge, insbesondere also in der Physik, mit den Axiomen der mathematischen Wahrscheinlichkeitsrechnung nicht auskommt, ondern eines besonderen Axioms bedarf; eines Axioms aber, das innerhalb nathematischer Begriffsbildung nicht sinnvoll ist, weil es aus der strengen Logik, der Logik mit zwei Wahrheitswerten, herausführt in eine Wahr-cheinlichkeitslogik, in der eine stetige Skala von Wahrscheinlichkeitswerten ür Aussagen besteht und Wahrheit und Falschheit nur noch Grenzfälle edeuten. Durch diese Wendung gelingt es, auch das Anwendungsproblem einer Lösung zuzuführen. Dabei werden zugleich die Ergebnisse unserer Axiomatik der mathematischen Wahrscheinlichkeitsrechnung wesentlich erden, denn es stellt sich heraus, daß mit dieser Axiomatik alle Pro-leme der Anwendung auf die Geltung eines einzigen Axioms zurück-eführt werden, dessen Anerkennung sich, wie ich glaube, niemand ent-iehen kann.

§ 2.

Die Wahrscheinlichkeitsimplikation.

Die Wahrscheinlichkeitsrechnung nimmt nur *einen* neuen Grundbegriff in sich auf; das ist der Begriff der *Wahrscheinlichkeitsimplikation*[4]).

Wenn wir von Wahrscheinlichkeit reden, so geschieht dies stets in der Form einer Beziehung zwischen Aussagen. Dem Auftreffen der Würfelseite 6 kommt eine Wahrscheinlichkeit nur dann zu, wenn gewürfelt wird; es liegt hier also eine Beziehung vor zwischen den beiden Aussagen „es wird gewürfelt" und „Seite 6 trifft auf". Man könnte hier auch von einer Beziehung zwischen Ereignissen reden; dem Ereignis $x =$ Würfelwurf und $y =$ Auftreffen der Seite 6; doch führt dies in logische Schwierigkeiten, weil dann die Einordnung der Wahrscheinlichkeitsimplikation unter die bekannten Beziehungen „Implikation", „und", „oder" usw., die auch Beziehungen zwischen Aussagen sind, unmöglich wird. Wir sprechen deshalb von einer Beziehung zwischen den Sätzen „x trifft ein" und „y trifft ein".

Doch bedarf es noch einer weiteren Bestimmung. Wenn wir die Wahrscheinlichkeitsimplikation ansetzen, so sind die vorkommenden Ereignisse stets derart charakterisiert, daß ihre Zugehörigkeit zu einer Klasse behauptet wird. Dies ist wesentlich für den Wahrscheinlichkeitsbegriff. Wenn etwa nach der Wahrscheinlichkeit für den tödlichen Ausgang einer Erkrankung gefragt wird, so wird man diese Angabe nur machen können, wenn man weiß, um welche Art von Krankheit es sich handelt; d. h. der Einzelfall muß erst in eine Klasse eingeordnet werden, ehe eine Wahrscheinlichkeit angegeben werden kann. Ordnet man denselben Einzelfall aber in eine andere Klasse ein, so ändert sich die Wahrscheinlichkeit. Ist etwa die Erkrankung eine schwere Grippe, so ergibt sich eine andere Todeswahrscheinlichkeit als für den Fall einer Grippe überhaupt. Dies ist kein Widerspruch; man hat nur daraus die Folgerung zu ziehen, daß die

---

[4]) Diesen Begriff habe ich zuerst benutzt in der Arbeit „Die Kausalstruktur der Welt und der Unterschied von Vergangenheit und Zukunft". Ber. d. Bayr. Akad. München 1925. Dort wurde (S. 147) auch bereits eine provisorische Tafel von Gesetzen der Wahrscheinlichkeitsimplikation aufgestellt; diese Tafel ist durch die im folgenden gegebene Axiomatik überholt. Für die weiteren Resultate jener Arbeit ist dies jedoch ohne Belang. — Die Auffassung der Wahrscheinlichkeit als einer Beziehung findet sich übrigens schon bei Bolzano (Wissenschaftslehre 1837, § 161); vgl. auch die Bemerkung von Dubislav in Erkenntnis I, 1930, S. 264; neuerdings hat auch Keynes (Über Wahrscheinlichkeit, Leipzig 1926, Barth) die Wahrscheinlichkeit als eine Beziehung behandelt und bereits, über Bolzano hinausgehend, einen Kalkül für diese Beziehung entwickelt. Doch gibt er dieser Beziehung eine so wesentlich andere und wie mir scheint, nicht haltbare Deutung, daß ich von seiner Arbeit für die vorliegende Untersuchung leider keinen Gebrauch machen konnte.

Wahrscheinlichkeit nicht einem Einzelfall, sondern einer Klasse zukommt. Dies wird auch daraus deutlich, daß die Wahrscheinlichkeitsaussage niemals auf Einzelereignisse beschränkt wird, sondern in gleicher Weise auf alle Ereignisse angewandt wird, die derselben Klasse angehören. Die Wahrscheinlichkeitsaussage ist deshalb eine generell behauptete Beziehung zwischen Aussagen über die Zugehörigkeit von Ereignissen zu Klassen.

Dies ist nicht immer deutlich, weil man es im täglichen Leben sehr oft scheinbar mit der Wahrscheinlichkeit von Einzelfällen zu tun hat. In dem Falle einer Erkrankung z. B. interessiert meist nur die eine bestimmte Person, und man möchte die Todeswahrscheinlichkeit nur für diese Person kennen. Man hilft sich hier, indem man den Fall in eine möglichst enge Klasse einschließt. Handelt es sich etwa um eine schwere Grippe, so wird man die für diese Klasse bestehende Todeswahrscheinlichkeit vorziehen der aus der allgemeineren Klasse der Grippekranken überhaupt berechneten Wahrscheinlichkeit. Gerade dieses Verfahren aber läßt den Klassencharakter der Wahrscheinlichkeit deutlich hervortreten. Man versucht nämlich den Einzelfall dadurch zu erfassen, daß man ihn als Grenzfall immer enger werdender Klassen darstellt, welche immer weitere Bestimmungsstücke des Einzelfalls berücksichtigen. Zugleich macht man dabei die Annahme, daß bei diesem Prozeß die Wahrscheinlichkeit der ineinander geschachtelten Klassen nach 1 bzw. 0 geht, daß man also mit immer größerer Sicherheit den Ausgang des vorliegenden Falles voraussehen kann. Dies ist nämlich die Behauptung des Kausalprinzips, wenn man dieses Prinzip von fiktiven Idealisierungen frei formuliert [5]). Aber an dem Klassencharakter der Wahrscheinlichkeitsaussage wird damit nichts geändert.

Nach dem Vorausgehenden haben wir der Wahrscheinlichkeitsaussage die Form zu erteilen: Wenn x ein Würfelwurf ist, so ist das darauffolgende Auftreffereignis y mit der Wahrscheinlichkeit $1/6$ ein Ereignis der Sechserklasse. Man bemerkt, daß hier noch eine Zuordnung zwischen x und y eingeht, dadurch gegeben, daß y dasjenige Auftreffereignis ist, welches durch den Wurf x ausgelöst wird. Eine derartige Zuordnung liegt bei allen Wahrscheinlichkeitsproblemen vor. Sie setzt voraus, daß die Klassen der Ereignisse x und y einander eineindeutig zugeordnet sind. Im allgemeinen handelt es sich hier um abzählbare Klassen; obgleich dies nicht notwendig

[5]) Ich habe schon in der oben genannten Arbeit (S. 138) darauf hingewiesen, daß die Annäherung der Wahrscheinlichkeit an 1 bzw. 0 auf diesem Wege nicht apriori gefordert werden kann, daß vielmehr mit der Möglichkeit einer von 1 bzw. 0 verschiedenen Grenze für die Wahrscheinlichkeit der Vorausberechnung gerechnet werden muß. Dieser Fall ist inzwischen von der Quantenmechanik aufgegriffen und als tatsächlich vorliegend erkannt worden. Vgl. auch meine Darstellung „Das Kausalproblem in der Physik", Naturwiss. 19 (1931), S. 713.

ist, wollen wir im folgenden doch der Einfachheit halber diese Annahme zugrunde legen[6]). Dann dürfen wir ohne Einschränkung der Allgemeinheit annehmen, daß jede dieser beiden Klassen in der Form einer numerierten Folge vorliegt, derart, daß die Zuordnung der Glieder zweier Folgen durch Gleichheit des Index gekennzeichnet ist.

Wir können nun für die Wahrscheinlichkeitsaussage die symbolische Form einführen. Wir benutzen für die Wahrscheinlichkeitsimplikation das Zeichen $\underset{p}{\rightarrow}$, die Zugehörigkeit des Elements zur Klasse bezeichnen wir in üblicher Weise mit dem Buchstaben ε, und als Allzeichen benutzen wir die übliche Einklammerung der Variablen. Damit schreiben wir:

$$(1) \qquad (i)\ (x_i \varepsilon O \underset{p}{\rightarrow} y_i \varepsilon P)$$

Von dieser Form sind alle Wahrscheinlichkeitssaussagen.

Es ist wesentlich, daß das Zeichen $\underset{p}{\rightarrow}$ stets hinter einem Allzeichen auftritt. Die Aussage

$$(2) \qquad x_i \varepsilon O \underset{p}{\rightarrow} y_i \varepsilon P$$

kommt niemals vor, und es ist deshalb nicht notwendig, ihr einen Sinn zuzuschreiben. Vielmehr hat für uns nur eine Aussage der Form (1) einen Sinn.

Die Formulierung (1) läßt die logische Stellung der Wahrscheinlichkeitsimplikation erkennen. Die Wahrscheinlichkeitsimplikation ist danach eine dreigliedrige Beziehung zwischen den Klassen O und P und dem Folgenpaar $x_i y_i$ (bzw. bei komplizierteren Ausdrücken einem Folgenmultipel). Dabei ist die Klasse P von der Klasse der y_i verschieden, und zwar wird es sowohl Elemente von P geben, die unter den y_i nicht vorkommen, als auch y_i, die nicht Elemente von P sind. Das gleiche gilt für die Klasse O und die Klasse der x_i, da wir uns nicht durchweg auf den Spezialfall beschränken können, wo wenigstens alle x_i auch Elemente von O sind. In der praktischen Anwendung wird häufig noch eine andere Deutung benutzt, in der an Stelle des Folgenpaares $x_i y_i$ eine Klasse derartiger Folgenpaare eintritt. Dies rührt daher, daß die zur Festlegung des Folgenpaares benutzten Vorschriften im allgemeinen sogleich eine Klasse derartiger Folgenpaare festlegen (z. B. legt die Vorschrift des Würfelns das Folgenpaar nicht nur für einen individuellen Würfel fest, sondern für alle

[6]) Für die von mir an anderer Stelle (Zeitschr. f. Phys. **53** (1929), S. 274) begründeten stetigen Wahrscheinlichkeitsfolgen ist diese vereinfachende Annahme aufzugeben und infolgedessen der weiterhin von uns benutzte Index i als stetige Veränderliche aufzufassen. Dies führt auf eine Änderung erst von der Stelle ab, wo die Häufigkeitsdeutung eingeführt wird; an Stelle der relativen Häufigkeit ist dann der Quotient zweier Maßgrößen zu setzen.

Würfel). Die Wahrscheinlichkeitsimplikation ist dann eine dreigliedrige Beziehung zwischen den Klassen O und P und der Klasse der Folgenpaare. Wir wollen vorerst die Wahrscheinlichkeitsimplikation nur in dem ersten, engeren Sinne auffassen; eine Übertragung unserer Formeln auf die Wahrscheinlichkeitsimplikation im weiteren Sinne ist ohne weiteres möglich.

Es sei noch eine Bemerkung über die Darstellung der Wahrscheinlichkeitsimplikation durch Satzfunktionen eingeschaltet. Jeder Ausdruck der Form $x_i \varepsilon O$ ist ja einer Satzfunktion $\varphi(x_i)$ äquivalent; entsprechend sei $\psi(y_i)$ die zu $y_i \varepsilon P$ äquivalente Satzfunktion. Um (1) schreiben zu können, müssen wir aber an Stelle der Paarliste $x_i y_i$ dann noch eine äquivalente Funktion $f(x_i, y_i)$ einführen. Dann nimmt (1) die Form an:

$$(3) \qquad (x_i)(y_i)[\varphi(x_i) \cdot f(x_i, y_i) \underset{p}{\Rightarrow} \psi(y_i)]$$

Auch in dieser zu (1) äquivalenten Schreibweise tritt also der Charakter der Wahrscheinlichkeitsimplikation als einer dreigliedrigen Beziehung hervor; ihre Glieder sind φ, f, ψ. Doch ist auch hier zu beachten, daß außer diesen Funktionen noch eine Ordnung ihrer Argumente im Sinne einer Folge gegeben sein muß, was wir durch die Einführung des Index i in (3) andeuten.

Die Schreibweise (1) erweist sich nun für weitere Zwecke als zu umständlich, und wir wollen jetzt eine abgekürzte Schreibweise einführen. Wir unterdrücken nämlich die Beziehung auf das Folgenpaar $x_i y_i$ und schreiben an Stelle von (1)

$$(4) \qquad (O \underset{p}{\Rightarrow} P)$$

Diese abgekürzte Schreibweise erhält eine anschauliche Deutung, wenn wir unter der *Folge* P (im Unterschied zur *Klasse* P) diejenige Teilfolge der y_i verstehen, für welche $y_i \varepsilon P$ gilt, und entsprechend von der Folge O usw. reden. Dann läßt sich die Wahrscheinlichkeitsbeziehung auch als zweigliedrige Beziehung zwischen den Folgen O und P auffassen. Allerdings unterdrücken wir in der Schreibweise (4) die Tatsache, daß die Folgen O und P in ganz bestimmter Weise einander zugeordnet sein müssen, wenn die Beziehung $\underset{p}{\Rightarrow}$ bestehen soll. Ein derartiges Verfahren ist erlaubt, wenn wir die Form (4) stets nur als eine Abkürzung für (1) auffassen und Regeln angeben, wie man einen Ausdruck der abgekürzten Schreibweise stets in die ausführliche Schreibweise übersetzen kann.

Daß wir mit der abgekürzten Schreibweise durchkommen, beruht auf der erwähnten Tatsache, daß das Zeichen $\underset{p}{\Rightarrow}$ nie ohne das vorangehende Allzeichen (i) auftritt. Wir können deshalb den runden Klammern in (4) die Bedeutung des Allzeichens zuweisen; die Variable i braucht hier nicht

genannt zu werden, weil sie stets in der ganzen Formel dieselbe ist. Hierauf beruht es, daß man für das Operieren mit den Formeln die Beziehung auf das Folgenpaar „vergessen" darf. Man darf sogar auch noch „vergessen", daß O und P Klassen sind, und von den „Ereignissen" O und P sprechen. Auf diese Weise gewinnt unser Kalkül eine ähnliche Form wie die klassische Wahrscheinlichkeitsrechnung.

Als Regel für den Übergang von der abgekürzten Schreibweise zur ausführlichen Schreibweise setzen wir nun fest: Man ersetzt überall einen großen Buchstaben A durch $x_i \varepsilon A$, wobei große Buchstaben, die sich nur durch den Index unterscheiden, dasselbe Individualzeichen x_i erhalten; außerdem setzt man (i) vor jede runde Klammer, welche große Buchstaben enthält. So lautet etwa der Ausdruck

$$(O \underset{r}{\rightarrow} [P_k \vee P_l] . Q) . (O \underset{p}{\rightarrow} P_k)$$

in der ausführlichen Schreibweise

$$(i)\,(x_i \varepsilon O \underset{r}{\rightarrow} [y_i \varepsilon P_k \vee y_i \varepsilon P_l] . [z_i \varepsilon Q]) . (i)\,(x_i \varepsilon O \underset{p}{\rightarrow} y_i \varepsilon P_k)$$

Wir müssen nun noch eine Festsetzung über die Negation treffen. Im Funktionenkalkül unterscheidet man zwischen den Aussagen $\overline{(i)\,\varphi(i)}$ und $(i)\,\overline{\varphi(i)}$; letztere ist die stärkere Form, weil aus ihr die erstere folgt, aber nicht umgekehrt. Wir setzen fest, daß die erste Form vorliegt, wenn der Negationsstrich sich über die runden Klammern mit ausdehnt, die zweite Form dagegen, wenn der Negationsstrich nur über den Zeichen innerhalb der Klammer steht. So gelten beispielsweise die folgenden Festsetzungen:

$$\overline{(O)} =_{Df} \overline{(i)\,(x_i \varepsilon O)} \qquad (\overline{O}) =_{Df} (i)\,\overline{(x_i \varepsilon O)}$$

Die Benutzung runder Klammern für das Allzeichen bewirkt, daß wir für andere Zusammenfassungen Klammern anderer Art, wie eckige Klammern, benutzen müssen. Im praktischen Rechnen brauchen wir auf die Sonderstellung der runden Klammern jedoch nur zu achten, wenn der Klammerausdruck negiert ist, da hier die genannte Unterscheidung eintritt; da jedoch in den praktisch wichtigen Formeln keine Negation auftritt, können wir den Unterschied runder und eckiger Klammern im allgemeinen außer acht lassen.

Im übrigen werden wir innerhalb unseres Kalküls die Formeln des Logikkalküls benutzen, und zwar dürfen wir, wie man leicht erkennt, für O, P, Q usw. die Regeln des Aussagenkalküls zugrunde legen, wenn die Umformungen innerhalb runder Klammern vollzogen werden, ebenso gelten die Regeln des Aussagenkalküls für Umformungen, bei welchen die runden Klammern als Ganzes auftreten. Für Rechnungen, welche die

nbestimmte Wahrscheinlichkeitsimplikation (7) enthalten, tritt der engere unktionenkalkül hinzu. Substitutionen, bei welchen Ausdrücke mit runden lammern innerhalb ebensolcher Ausdrücke eingeführt werden, lassen wir orerst nicht zu. Insbesondere werden die folgenden Formeln des Logikalküls gebraucht:

$$(5a)\ a \vee \bar{a} \qquad (5b)\ a \supset b \equiv \bar{a} \vee b \qquad (5c)\ a.(b \vee c) \equiv a.b \vee a.c$$

$$(5d)\ a.b \equiv b.a \qquad (5e)\ a \vee b \equiv b \vee a \qquad (5f)\ \bar{\bar{a}} \equiv a \qquad (5g)\ \overline{a.b} \equiv \bar{a} \vee \bar{b}$$

$$(5h)\ \overline{(x)\varphi(x)} \equiv (\exists x)\overline{\varphi(x)} \qquad (5i)\ \overline{(\exists x)\varphi(x)} \equiv (x)\overline{\varphi(x)}$$

$$(5k)\ \varphi(y) \supset (\exists x)\varphi(x) \qquad (5l)\ (\exists x)\varphi(x) \supset \varphi(\varepsilon_x \varphi(x))\ ^{6a})$$

Wir wählen hier die Russellschen Zeichen [7]) mit Ausnahme der Negation, ie wir nach Hilbert durch Überstreichung kennzeichnen. Unser Zeichen ir die Wahrscheinlichkeitsimplikation ist in Anlehnung an das Russellsche mplikationszeichen gebildet; auch setzen wir das bedingende Glied als orderglied, gerade wie bei der strengen Implikation. Der entgegenstehende prachgebrauch „P folgt aus O", „P besitzt die Wahrscheinlichkeit p in ezug auf O" hat sich für beide Implikationen als unzweckmäßig erwiesen.

Für manche Zwecke bedürfen wir eines Symbols, um den Ausdruck (3) ach p aufzulösen. Wir führen deshalb die Schreibweise ein

$$6) \qquad W(O, P) = p$$

as wir lesen können: „die Wahrscheinlichkeit von O nach P ist gleich p" a das Symbol $W(\ )$ niemals allein auftritt, sondern immer nur in Gleiungen der Form (6), brauchen wir nur dieser Form eine Bedeutung zu teilen; wir setzen fest, daß (6) dasselbe bedeutet wie (4).

Eine weitere Abkürzung ist gegeben durch die folgende Festsetzung

$$7) \qquad (O \underset{z}{\succ} P) =_{Df} (\exists z).(O \underset{z}{\succ} P) =_{Df} (\exists z)(i)(x_i \varepsilon O \underset{z}{\succ} y_i \varepsilon P)$$

6a) Dies ist das Hilbertsche ε-Axiom (Grundlagen der Mathematik, Abh. d. math. m. Hamburg 1928); wir werden es in § 4 brauchen, um für eine durch Existentialichen gebundene Variable Bestimmungen treffen zu können.

7) $a.b$ = „a und b", $a \vee b = a$ oder b", $a \supset b$ = „a impliziert b", $a \equiv b$ = „a äquilent b" (Implikation nach beiden Richtungen), (x) = „für alle x", $(\exists x)$ = „es gibt x", ε = „ist Element von". Wendet man die Formeln (5a) bis (5g) auf unsere rößen O, P, Q an, so sind in strenger Betrachtung die gebrauchten Zeichen übrigens cht die Zeichen des Aussagenkalküls, sondern die entsprechenden Zeichen des esem isomorphen Klassenkalküls, da O, P, Q Klassen sind; erst bei Rückgang auf e ausführliche Schreibweise (1) treten die entsprechenden Zeichen des Aussagenlküls auf. Für das Rechnen mit den gekürzt geschriebenen Formeln ist diese Unterheidung jedoch unwesentlich. Zur Ersparung von Klammern setzen wir nun fest, ß die Implikation, ebenso auch die Wahrscheinlichkeitsimplikation, das stärker ennende Zeichen ist gegenüber „und" und „oder" und „ist Element von".

Das Zeichen $\overset{p}{\rightarrow}$ nennen wir unbestimmte Wahrscheinlichkeitsimplikation. — Es sei noch bemerkt, daß entsprechend einer allgemeinen Regel des Funktionenkalküls stets gilt

$$(8) \qquad (O \overset{p}{\rightarrow} P) \equiv (\exists x)(O \overset{x}{\rightarrow} P).(x = p)$$

Mit den entwickelten Zeichen können wir jetzt die Axiomatik der Wahrscheinlichkeitsrechnung durchführen. Doch sei noch eine Vorbemerkung vorausgeschickt. Wir betrachten die Entscheidung darüber, ob für gegebene O und P eine bestimmte Wahrscheinlichkeitsimplikation nach (4) besteht, als eine Tatsachenfrage, deren Beantwortung der Empirie angehört, nicht aber der mathematischen Wahrscheinlichkeitsrechnung. Diese hat es vielmehr nur mit Umformungen derartiger Ausdrücke zu tun. Die Frage, wie die empirische Ermittlung erfolgt, wird uns erst in § 11 beschäftigen. Wir betrachten es auch schon als eine empirische Feststellung, ob zwischen O und P eine unbestimmte Wahrscheinlichkeitsimplikation nach (7) besteht. Von manchen wird die Ansicht vertreten, daß dies für alle beliebigen Ereignisse gelte. Wir möchten jedoch diese Annahme vorerst nicht einführen. (Vgl. den Schluß von § 3.)

Sodann aber treffen wir vorerst auch keinerlei Entscheidung darüber, was der Sinn der Wahrscheinlichkeitsaussage ist. Wir lassen es also zunächst ganz offen, was mit dem Bestehen der Wahrscheinlichkeitsimplikation und insbesondere der Angabe eines Wahrscheinlichkeitsgrades gemeint ist. Wir stellen vielmehr zunächst nur das System der Gesetze der Wahrscheinlichkeitsimplikation auf; die so entwickelte Wahrscheinlichkeitsrechnung nennen wir *formale Wahrscheinlichkeitsrechnung*. Wir benutzen also zunächst auch noch keinerlei Häufigkeitsbegriffe. Infolgedessen machen wir vorerst auch noch keinen Gebrauch von der Ordnung der Folgen x_i und y_i; vielmehr gehen diese in die formale Wahrscheinlichkeitsrechnung nur als *gegenseitig zugeordnete* Klassen, nicht als *in sich geordnete* Klassen, ein. Wir haben aber die Forderung der Ordnung in unsere Begriffsbildung bereits aufgenommen, weil wir sie für die später einzuführende Häufigkeitsdeutung der Wahrscheinlichkeit brauchen.

Irgendeine Begründung oder Deutung des Wahrscheinlichkeitsgrades aus logischen Überlegungen, wie dies z. B. in dem Begriff der gleichmöglichen Fälle oder durch das Prinzip des mangelnden Grundes versucht wird, lehnen wir also als vollständig unhaltbar ab. In dieser Auffassung, die wir schon immer vertreten haben[8]), treffen wir uns mit der Auffassung von Herrn v. Mises[9]); weitere Übereinstimmungen mit Herrn v. Mises liegen darin, daß

[8]) H. Reichenbach, Der Begriff der Wahrscheinlichkeit für die mathematische Darstellung der Wirklichkeit. Dissertation, Erlangen 1915.

[9]) R. v. Mises, Grundlagen der Wahrscheinlichkeitsrechnung, Math. Zeitschr. 5 (1919), S. 62.

vir die Wahrscheinlichkeitsimplikation als eine nur für Folgen, nicht für Einzelereignisse definierte Beziehung ansehen, ferner in der von uns allerdings erst an späterer Stelle eingeführten Häufigkeitsdeutung. Nachdem wir n einigen andern Punkten (§ 7) andern Auffassungen von Herrn v. Mises widersprechen müssen, seien diese Übereinstimmungen hier sehr gern vorangestellt.

Gegenüber der klassischen Wahrscheinlichkeitsrechnung, mit der unser ormaler Kalkül gewisse Ähnlichkeiten aufweist, besitzt unsere Auffassung den tiefgehenden Unterschied, daß die Wahrscheinlichkeit von vornherein als *Beziehung* eingeführt wird. Die klassische Wahrscheinlichkeitsrechnung behandelt die Wahrscheinlichkeit als *Eigenschaft*, d. h. als Funktion nur einer Veränderlichen; hieraus sind der klassischen Wahrscheinlichkeitsrechnung manche Schwierigkeiten erwachsen. Die Auffassung der Wahrscheinlichkeit als Beziehung, d. h. als Funktion mehrerer Veränderlicher, hat den großen Vorzug, daß die bekannte Unterscheidung von „Absolutwahrscheinlichkeit" und „Relativwahrscheinlichkeit" überflüssig wird. Die Wahrscheinlichkeit ist stets der Grad einer Relation, und es ist kein tieferer Unterschied, ob das Vorderglied aus *einem* Zeichen oder aus einer Konjunktion von Zeichen besteht. Weiter werden wir auch den Begriff der „apriori-Wahrscheinlichkeit" vermeiden können.

Aber auch im Vergleich zu der Wahrscheinlichkeitsrechnung von Herrn v. Mises ist unsere Wahrscheinlichkeitsrechnung wesentlich allgemeiner. So schon deshalb, weil wir die Wahrscheinlichkeit abhängiger Ereignisse einbegreifen, also die Wahrscheinlichkeitsrechnung des allgemeinen Multiplikationstheorems entwickeln und uns nicht auf das spezielle Multiplikationstheorem beschränken. Weiterhin deshalb, weil wir die Wahrscheinlichkeit von vornherein als eine Beziehung ansetzen, die Folgenmultipel enthält, während für Herrn v. Mises die Wahrscheinlichkeit notwendig an eine einzige Folge gebunden ist. Für unsere Auffassung stellt sich nämlich das v. Misessche Kollektiv als der Spezialfall dar, der entsteht, wenn die Zuordnung der x_i zu den y_i zur Identität wird; das v. Misessche Kollektiv nimmt also bei uns die Form an

$$(i)(y_i \varepsilon P \underset{p}{\rightarrow} y_i \varepsilon P_k)$$

oder abgekürzt

$$(P \underset{p}{\rightarrow} P_k) \tag{9}$$

In der Darstellung nach (3) wird dann die Funktion f entbehrlich. Die Beschränkung auf diesen Spezialfall, der für manche Anwendungen, z. B. in Sozialstatistiken, in Betracht kommt, läßt jedoch manche Eigenschaften des Wahrscheinlichkeitsbegriffes unerkennbar werden, insbesondere die Anwendung der Wahrscheinlichkeitsimplikation zur Deutung des Begriffs der

Kausalität. Aber auch für den Kalkül selbst bringt unsere Auffassung Vorzüge mit. In der v. Misesschen Wahrscheinlichkeitsrechnung werden durch Operationen aus bestehenden Folgen neue Folgen gebildet, für welche die Wahrscheinlichkeit dann nach (9) definiert wird; auf diese Weise werden aus bestehenden Wahrscheinlichkeiten neue Wahrscheinlichkeiten berechnet. Bei unserm Kalkül dagegen läßt sich die neue Wahrscheinlichkeit durch eine Beziehung auf ein Folgenmultipel in gleicher Weise deuten, wie die ursprüngliche Wahrscheinlichkeit durch Beziehung auf ein Folgenmultipel gedeutet wird. Diese formale Vereinfachung ermöglicht den Gebrauch der einfachen Operationen „und" und „oder" an Stelle der v. Misesschen Operationen der „Verbindung", „Mischung" usw.

Zur Methodik der folgenden Untersuchungen sei noch bemerkt, daß wir uns der Einfachheit halber auf Folgen mit Alternativcharakter, die also nur die Ereignisse P und $\bar{P}$ unterscheiden, beschränken werden. Hierin liegt keine Beschränkung der Allgemeinheit. Wie Herr v. Mises gezeigt hat, kann man jede Behandlung von Folgen mit stetig veränderlichem Merkmal auf die Behandlung von Alternativfolgen zurückführen, indem man verlangt, daß jede beliebige Teilung des Merkmalraums in zwei Gebiete eine Alternativfolge der geforderten Eigenschaften erzeugt.

§ 3.

Die Axiome der Wahrscheinlichkeitsrechnung.

I. *Eindeutigkeit*:

$$(p \dotplus q) \succ [(O \underset{p}{\rightarrow} P).(O \underset{q}{\rightarrow} P) \equiv (\bar{O})]$$

II. *Normierung*:

1. $(O \succ P) \succ (O \underset{p}{\rightarrow} P).(p = 1)$
2. $(O \underset{p}{\rightarrow} P) \succ (p \geqq 0)$

III. *Additionstheorem*:

1. $(O \underset{p}{\rightarrow} P).(O \underset{q}{\rightarrow} Q).(O.P \succ \bar{Q}) \succ (O \underset{r}{\rightarrow} P \vee Q).(r = p + q)$
2. $(O \underset{p}{\rightarrow} P).(O \underset{r}{\rightarrow} P \vee Q).(O.P \succ \bar{Q}) \succ (O \rightarrow Q)$

IV. *Multiplikationstheorem*:

1. $(O \underset{p}{\rightarrow} P).(O.P \underset{u}{\rightarrow} Q) \succ (O \underset{w}{\rightarrow} P.Q).(w = p \cdot u)$
2. $(O \underset{p}{\rightarrow} P).(O \underset{w}{\rightarrow} P.Q) \succ (O.P \rightarrow Q)$
3. $(O.P \underset{u}{\rightarrow} Q).(O \underset{w}{\rightarrow} P.Q) \succ (O \rightarrow P)$

Zu diesen Axiomen treten, nach einer vorangehenden Bemerkung, die Axiome des Aussagen- und engeren Funktionenkalküls noch hinzu. In bezug auf die Schreibweise sei noch bemerkt, daß die Hinzufügung des Wahrscheinlichkeitsgrades in einer nachgesetzten Klammer bei komplizierteren Ausdrücken, wie auf der rechten Seite von Axiom III, 1, lediglich zur besseren Übersichtlichkeit erfolgt; ihre Rechtfertigung findet diese Schreibweise durch (8, § 2), so daß in derartigen Formeln stets ein Existentialoperator vor die runde Klammer gesetzt werden darf.

Zu unserer Fassung der Axiome ist noch folgendes zu bemerken. Neu gegenüber bisherigen Darstellungen sind zunächst die Axiome I und II, 1. II, 1 formuliert den Zusammenhang von Gewißheit und Wahrscheinlichkeit 1; es ist bemerkenswert, daß hier die Implikation nur in *einer* Richtung gilt. Was gewiß ist, ist auch mit der Wahrscheinlichkeit 1 zu erwarten; aber das Umgekehrte gilt nicht notwendig[10]). I legt die Eindeutigkeit des Grades der Wahrscheinlichkeitsimplikation fest, und zwar in folgendem Sinne. Aus I folgt durch bekannte logische Umformungen, in denen das Äquivalenzzeichen $\equiv$ durch die Implikation nach beiden Richtungen, die Implikation aber nach dem Schema (5b, § 2) aufgelöst wird:

$$(1) \qquad \overline{(\bar{O})} \,.\, (O \underset{p}{\rightarrow} P) \,.\, (O \underset{q}{\rightarrow} P) \supset (p = q)$$

Nun ist aber nach (5h, § 2) die Bedeutung des ersten hier auftretenden Zeichens in der ausführlichen Schreibweise gegeben durch

$$(2) \qquad \overline{(\bar{O})} \equiv (\exists i) \,.\, (x_i \,\varepsilon\, O)$$

Danach bedeutet (1), daß immer dann, wenn in einer Folge überhaupt Glieder $x_i \,\varepsilon\, O$ auftreten, die Wahrscheinlichkeitsimplikation eindeutig ist. Mehr braucht nicht verlangt zu werden. Vielmehr besagt I, daß diese Eindeutigkeit nicht mehr besteht, wenn die Folge überhaupt keine Glieder $x_i \,\varepsilon\, O$ enthält. Dieser durch $(\bar{O})$ charakterisierte Fall führt nämlich wegen I, wenn man hier die von rechts nach links führende Implikation benutzt, auf

$$(3) \qquad (\bar{O}) \supset (O \underset{p}{\rightarrow} P)$$

Da hier die auf der rechten Seite stehenden Zeichen P und p nicht durch links stehende Zeichen bestimmt sind, unsere Formeln aber für jede beliebige Substitution von Zeichen Geltung behalten sollen, besagt dies, daß man im Falle $(\bar{O})$ jede beliebige Wahrscheinlichkeitsimplikation, für beliebige P und mit beliebigem Grad p, an O anschließen kann. Zu dem

[10]) Man erkennt dies am besten aus der Häufigkeitsdeutung, weil die Wahrscheinlichkeit 1 nur verlangt, daß die Häufigkeit im limes nach 1 geht, während einzelne Fälle sehr wohl vorkommen können, wo P auf O folgt. Für das Folgende machen wir jedoch von der Häufigkeitsdeutung zunächst keinen Gebrauch.

Verzicht auf die Eindeutigkeit für diesen Fall zwingt uns eine entsprechende Eigenschaft der strengen Implikation. Da nämlich eine unwahre Prämisse jeden Satz impliziert, gilt

$$(4) \qquad (\bar{O}) \succ (O \succ P).(O \succ \bar{P})$$

und dies führt wegen II, 1 und Satz 6, § 4, auf Mehrdeutigkeit des Wahrscheinlichkeitsgrades für den Grenzfall 0 bzw. 1. In I übertragen wir diese Mehrdeutigkeit unter Erweiterung für alle Grade sinngemäß auf die Wahrscheinlichkeitsimplikation. Übrigens ist I nichts als eine erweiterte Nachbildung der logischen Formel

$$(5) \qquad [O \succ P].[O \succ \bar{P}] \equiv \bar{O}$$

die auch in der Form

$$(6) \qquad (O \succ P).(O \succ \bar{P}) \equiv (\bar{O})$$

noch gilt. Während aber (5) im Aussagenkalkül aus anderen Axiomen ableitbar ist, müssen wir die entsprechende Formel I in der Wahrscheinlichkeitsrechnung als Axiom aufstellen, weil für die Wahrscheinlichkeitsimplikation eine (5b, § 2) entsprechende Auflösung nicht besteht.

IV, 1 ist in seiner Fassung abweichend von den üblichen Darstellungen des Multiplikationstheorems. Es läßt sich zeigen, daß dieses Theorem nur in der hier gegebenen Fassung immer wahr ist; der Grund hierfür wird allerdings erst in § 6 hervortreten[11]). — Andere Fassungen sind als Spezialfälle anzusehen, deren Geltung an bestimmte Bedingungen gebunden ist; wir nennen IV, 1 deshalb *allgemeines Multiplikationstheorem*. Einige derartige Spezialisierungen seien hier kurz dargestellt.

Die erste Spezialisierung entsteht, wenn

$$(7) \qquad W(O.P, Q) = W(O, Q)$$

IV, 1 geht dann in das *spezielle Multiplikationstheorem* über:

$$(8) \qquad (O \underset{p}{\succ} P).(O \underset{q}{\succ} Q) \succ (O \underset{w}{\succ} P.Q).(w = p \cdot q)$$

Wir sagen dann: „die Ereignisse P und Q sind unabhängig voneinander in bezug auf O". Die Unabhängigkeit ist eine dreigliedrige Beziehung, nicht eine zweigliedrige; man kann Fälle konstruieren, wo P und Q unabhängig voneinander in bezug auf O sind, aber nicht in bezug auf ein anderes O'.

---

[11]) In der englisch-amerikanischen Literatur dagegen wird diese Fassung schon seit langem benutzt; so wird sie schon 1878 von Ch. Peirce angegeben (vgl. die Neuausgabe seiner Schriften: Ch. Peirce, Chance, Love, and Logic, ed. by M. Cohen, New York 1923, Harcourt, S. 83). Auch Keynes benutzt diese Fassung, vgl. [4]) S. 95. Doch scheint es, daß die Begründung dieser Fassung, die wir in § 6 geben, dort noch nicht deutlich erkannt worden ist.

aß die Unabhängigkeitsbeziehung in bezug auf P und Q symmetrisch ist. lgt erst aus Satz 9a, § 4; denn aus diesem Satz folgt, daß mit (7) auch $(O.Q,P)=W(O,P)$ ist.

Eine andere Spezialisierung von IV, 1 erhält man, wenn O sich als Konunktion zweier Ereignisse A und B darstellen läßt, derart, daß

9) $$W(A.B,P)=W(A,P), \qquad W(A.B.P,Q)=W(B.P,Q)$$

V, 1 führt dann auf

10) $$(A \underset{p}{\rightarrow} P).(B.P \underset{u}{\rightarrow} Q) \succ (A.B \underset{w}{\rightarrow} P.Q).(w=p\cdot u)$$

Kommt hierzu noch die zu (7) analoge Spezialisierung

11) $$W(B.P,Q)=W(B,Q)$$

o entsteht

12) $$(A \underset{p}{\rightarrow} P).(B \underset{q}{\rightarrow} Q) \succ (A.B \underset{w}{\rightarrow} P.Q).(w=p\cdot q)$$

Beispiel hierfür ist das Werfen mit zwei Würfeln, A bedeutet dann den Wurf mit dem einen Würfel, B den Wurf mit dem andern. Ohne die Bedingungen (9) und (11) ist (12) jedoch nicht zulässig.

Eine weitere Spezialisierung entsteht, wenn

13) $$W(O.P,Q)=W(P,Q)$$

Dann wird aus IV, 1:

14) $$(O \underset{p}{\rightarrow} P).(P \underset{u}{\rightarrow} Q) \succ (O \underset{w}{\rightarrow} P.Q).(w=p\cdot u)$$

Beispiele hierfür treten in Kausalverkettungen auf; O sei etwa das Fallen des Barometers, P das Eintreten von Regen, Q das Eintreten einer Überschwemmung. Doch ist bei Anwendung von (14) jedesmal zu prüfen, ob (13) erfüllt ist. — Der Schluß auf $(O \underset{q}{\rightarrow} Q)$ und die Berechnung von q kommt übrigens auch hier erst mit Satz 10, § 4 zustande.

Neu gegenüber den üblichen Darstellungen sind ferner die Umkehraxiome III, 2 und IV, 2 bis 3. Steht man auf dem Standpunkt, daß zwischen irgend zwei Ereignissen stets eine Wahrscheinlichkeitsimplikation besteht, so sind diese Axiome überflüssig, da sie rechts nur die unbestimmte Wahrscheinlichkeitsimplikation enthalten, also nur das Bestehen einer Wahrscheinlichkeitsimplikation behaupten, während der Grad der Wahrscheinlichkeitsimplikation schon durch III, 1 bzw. IV, 1 bestimmt ist (Satz 1, 6, 7, § 4). Da wir diesen Standpunkt nicht teilen, mußten wir diese Axiome aufnehmen, weil sie bei den Sätzen benutzt werden. Übrigens müßte bei Fortlassung dieser Axiome ein anderes Axiom aufgenommen werden, welches das Bestehen der unbestimmten Wahrscheinlichkeitsimplikation für irgend zwei Ereignisse postuliert.

§ 4.

Ableitbare Sätze.

Wir wollen jetzt die wichtigeren Sätze zusammenstellen, die sich aus dem Axiomensystem ableiten lassen. Einzelne dieser Sätze werden besonders übersichtlich, wenn man sie in der *W*-Schreibweise wiedergibt; wir werden in solchen Fällen diese Schreibweise hinzufügen und die betreffende Formel mit der gleichen Nummer, unter Hinzufügung des Buchstabens a, bezeichnen. Die korrekte Darstellung ist jedoch nur durch die implikative Schreibweise gegeben, da wir in der *W*-Schreibweise, der Übersichtlichkeit halber, vielfach die Bedingungen weglassen werden, unter denen der Satz gilt; diese Bedingungen sind dann in Worten hinzugefügt. Die *W*-Schreibweise soll nur den numerischen Teil der Formel wiedergeben.

Unter den abzuleitenden Sätzen werden sich auch einige Ungleichungen befinden, denen die vorzugebenden Wahrscheinlichkeitsgrade genügen müssen Es handelt sich hier um einschränkende Bedingungen, vergleichbar den einschränkenden Bedingungen, welche in der Geometrie für die Konstruktion eines Dreiecks aus vorzugebenden Größen bestehen. Zu diesen Bedingungen rechnen wir auch die Forderung, daß der Nenner eines Wahrscheinlichkeitsgrades niemals 0 werden darf; diese Forderung soll im folgenden nicht weiter hingeschrieben werden, sondern als selbstverständliche Verabredung gelten. Sie gilt insbesondere auch für die Axiome IV, 2—3, wie sich aus Satz **2** und **3** ergibt.

Wir wollen zunächst zeigen, daß in den Umkehraxiomen die unbestimmte Wahrscheinlichkeitsimplikation durch eine bestimmte Wahrscheinlichkeitsimplikation im Sinne des ersten Axioms der betreffenden Gruppe ersetzt werden darf. Für diesen Beweis ist die in I geforderte Eindeutigkeit wesentlich. Bei dem Beweis müssen wir Existentialzeichen einführen, da wir die Definition (7, § 2) der unbestimmten Wahrscheinlichkeitsimplikation einsetzen müssen; dies ist jedoch der einzige derartige Fall, weiterhin werden wir an Stelle der Umkehraxiome die ihnen entsprechenden Sätze 1—3 benutzen und dann für die weiteren Formeln keine Existentialzeichen mehr benötigen. Auch Axiom I werden wir von da ab nicht mehr besonders benutzen; dieses Axiom kommt für die hier abzuleitenden Sätze nur auf dem Weg über die Sätze 1—3 in Betracht. Dadurch werden die späteren Rechnungen wesentlich einfacher; sie können ohne Benutzung der in I auftretenden Negation innerhalb des Klammerzeichens auskommen.

1. $$(O \underset{p}{\rightarrow} P).(O \underset{r}{\rightarrow} P \vee Q).(O.P > \bar{Q}) > (O \underset{q}{\rightarrow} Q).(q = r - p)$$

Der Beweis dieses zunächst unmittelbar einleuchtenden Satzes ist verhältnismäßig kompliziert und sei der Ordnung halber hier in aller Strenge angegeben.

Für diesen Beweis sei die linke Seite von Satz 1 mit a bezeichnet, ferner benutzen wir die Abkürzungen:

$$(O \underset{x}{\rightarrow} Q) = \varphi(x) \qquad [x = r - p] = \psi(x)$$

Nach (8, § 2) bzw. einer zu Anfang von § 3 gemachten Bemerkung können wir Satz 1 dann schreiben:

$$a \supset (\exists x)\, \varphi(x) . \psi(x)$$

Nun gilt wegen (5k und 5l, § 2)

$$(\exists x)\, \varphi(x) \equiv \varphi(\varepsilon_x \varphi(x)) \equiv \varphi(\varepsilon)$$

wenn zur Abkürzung

$$\varepsilon = \varepsilon_x \varphi(x)$$

gesetzt wird. Sodann benutzen wir die Abkürzung

$$A = a . \varphi(\varepsilon) . \overline{\psi(\varepsilon)}$$

Nun enthält A u. a. die Ausdrücke

$$(O \underset{p}{\rightarrow} P) . (O \underset{r}{\rightarrow} Q) . (O . P \supset Q)$$

also folgt nach III, 1:

$$A \supset (O \underset{s}{\rightarrow} P \vee Q) . (\varepsilon = p + \varepsilon)$$

Nehmen wir noch die Bedeutung von $\overline{\psi(\varepsilon)}$ hinzu, so folgt auch $s \neq r$; es gilt deshalb, da A über a auch $(O \underset{r}{\rightarrow} P \vee Q)$ enthält:

$$A \supset (O \underset{r}{\rightarrow} P \vee Q) . (O \underset{s}{\rightarrow} P \vee Q) . (s \neq r)$$

Wenden wir hierauf Axiom I an, so folgt, wenn wir zur Abkürzung $(\overline{O}) = b$ setzen und für A seine Bedeutung einsetzen:

$$a . \varphi(\varepsilon) . \overline{\psi(\varepsilon)} \supset b$$

Hieraus folgt durch einfache Umrechnung unter Benutzung von (5b, 5f, 5g, § 2)

$$a . \bar{b} \supset \overline{\varphi(\varepsilon) . \overline{\psi(\varepsilon)}} \supset [\varphi(\varepsilon) \supset \psi(\varepsilon)]$$

Da außerdem nach Axiom III, 2 auch gilt

$$a \supset (\exists x)\, \varphi(x) \supset \varphi(\varepsilon)$$

so gilt auch

$$a . \bar{b} \supset \varphi(\varepsilon) . [\varphi(\varepsilon) \supset \psi(\varepsilon)] \supset \varphi(\varepsilon) . \psi(\varepsilon)$$

Nun gilt auch, wenn man jetzt $\varphi(\varepsilon) . \psi(\varepsilon)$ als *eine* Funktion $\chi(\varepsilon)$ auffaßt, wegen (5k, § 2)

$$\varphi(\varepsilon) . \psi(\varepsilon) \supset (\exists x)\, \varphi(x)\, \psi(x)$$

also gilt auch

$$a . \bar{b} \supset (\exists x)\, \varphi(x) . \psi(x)$$

Andererseits folgt aber auch aus Axiom I, wenn man die aus diesem Axiom abgeleitete Formel (3, § 3) zugrunde legt, unter Berücksichtigung der Bedeutung von $\varphi(x)$ und $\psi(x)$

$$a . b \supset (\exists x)\, \varphi(x) . \psi(x)$$

Aus den beiden letzten Formeln ergibt sich durch leichte Umrechnung

$$a \supset (\exists x)\, \varphi(x) . \psi(x)$$

Dies ist aber Satz 1.

Man kann III, 1 und Satz **1** in der Formel der W-Schreibweise zusammenfassen:

1a. $$W(O, P \vee Q) = W(O, P) + W(O, Q)$$

Diese Formel gilt nur für einander ausschließende Ereignisse P und Q. — Daß hier zwei Formeln der implikativen Schreibweise durch nur eine Formel der W-Schreibweise ausgedrückt werden, liegt daran, daß in der Gleichungsschreibweise die Richtung der logischen Implikation nicht zum Ausdruck kommt; man kann hier irgendwelche Größen als gegeben betrachten und danach die anderen berechnen. Ihre Rechtfertigung findet diese Schreibweise deshalb erst durch die Aufstellung der zwei Axiome III, 1 und III, 2, welche postulieren, daß jede der drei Größen durch die beiden andern bestimmt ist.

2. $$(O \underset{p}{\rightarrow} P) . (O \underset{w}{\rightarrow} P.Q) > (O.P \underset{u}{\rightarrow} Q) . \left(u = \frac{w}{p}\right)$$

3. $$(O.P \underset{u}{\rightarrow} Q) . (O \underset{w}{\rightarrow} P.Q) > (O \underset{p}{\rightarrow} P) . \left(p = \frac{w}{u}\right)$$

Der Beweis dieser beiden Sätze erfolgt aus IV, 1 bis 3 und I durch entsprechende Schlüsse wie zu Satz 1.

Man kann IV, 1 zusammen mit Satz **2** und **3** in der Formel wiedergeben:

3a. $$W(O, P.Q) = W(O, P) \cdot W(O.P, Q)$$

Wegen der Möglichkeit dieser Zusammenfassung vgl. die Bemerkung zu **1a.**

4. $$(O \underset{p}{\rightarrow} P) > (O \underset{u}{\rightarrow} \bar{P}) \cdot (u = 1 - p)$$

4a. $$W(O, P) + W(O, \bar{P}) = 1$$

Zum Beweis nehmen wir zu $(O \underset{p}{\rightarrow} P)$ den immer wahren Satz $(O > P \vee \bar{P})$ hinzu, den wir nach II, 1 umformen; beides in Satz 1 eingesetzt, ergibt **4.** **4** nennen wir auch *Ergänzungssatz.*

5. $$(O \underset{p}{\rightarrow} P) > (0 \leqq p \leqq 1)$$

5a. $$0 \leqq W(O, P) \leqq 1$$

Beweis aus II, 2 und Satz **4.**

6. $$(O > \bar{P}) > (O \underset{p}{\rightarrow} P) . (p = 0)$$

Beweis aus II, 1 und Satz **4.**

7. $$(O \underset{u}{\rightarrow} P.Q) . (O \underset{v}{\rightarrow} P.\bar{Q}) > (O \underset{p}{\rightarrow} P) . (p = u + v)$$

7a. $$W(O, P) = W(O, P.Q) + W(O, P.\bar{Q})$$

Beweis aus III, 1 mit $P.Q \vee P.\bar{Q} = P.[Q \vee \bar{Q}] = P.$

Auch **7a** läßt sich als Gleichung auffassen, wie früher **1a**, da man die entsprechenden Umkehrungen zu 7 leicht beweisen kann.

8. $(O \underset{p}{\rightarrow} P).(O \underset{q}{\rightarrow} Q).(O \underset{w}{\rightarrow} P.Q) > (O \underset{r}{\rightarrow} P \vee Q).(r = p + q - w)$

8a. $W(O, P \vee Q) = W(O, P) + W(O, Q) - W(O, P.Q)$

Zum Beweis schließt man zunächst mit der Umkehr von Satz 7 (entsprechend **7a**) auf $(O \underset{s}{\rightarrow} P.\bar{Q}).(s = p - w)$, ebenso auf $(O \underset{t}{\rightarrow} \bar{P}.Q).(t = q - w)$. Mit III, 1 erhält man dann 8, wenn man beachtet, daß

$$P \vee Q = P.Q \vee P.\bar{Q} \vee \bar{P}.Q$$

Die Rechnung wird in der W-Schreibweise wesentlich übersichtlicher:

Aus **1a**: $W(O, P \vee Q) = W(O, P.Q) + W(O, \bar{P}.Q) + W(O, P.\bar{Q})$

Aus **7a**: $0 = W(O, P.Q) - W(O, P) \quad + W(O, P.\bar{Q})$

Aus **7a**: $0 = W(O, P.Q) + W(O, \bar{P}.Q) - W(O, Q)$

Für die Gewinnung der letzten Gleichung ist in **7a** P und Q vertauscht. Subtrahiert man die beiden letzten Gleichungen von der ersten, so entsteht **8a**. — Wir nennen 8 *allgemeines Additionstheorem*, im Gegensatz zu dem *speziellen Additionstheorem* III, 1, das aus Satz 8 für den Spezialfall $W(O, P.Q) = 0$ entsteht. Der Wert von 8 besteht darin, daß diese Formel die oder-Wahrscheinlichkeit angibt, ohne an die Beschränkung auf sich ausschließende Ereignisse gebunden zu sein. Deshalb kann auch die Schreibweise **8a** ohne Einschränkung benutzt werden, im Gegensatz zu **1a**.

Aus **8a** leiten wir noch eine Ungleichung ab. Da wegen Satz 5 und **3a** $w = q \cdot v \leqq q$, folgt $p \leqq p + q - w$. Wir schreiben dies

8b. $W(O, P \vee Q) \geqq W(O, P)$

Dies ist eine einschränkende Bedingung für die Wahl vorzugebender Größen.

9. $(O \underset{p}{\rightarrow} P).(O.P \underset{u}{\rightarrow} Q).(O \underset{q}{\rightarrow} Q) > (O.Q \underset{v}{\rightarrow} P).\left(v = \frac{p \cdot u}{q}\right)$

Beweis aus IV, 1 und Satz 2, wenn in letzterem P mit Q vertauscht wird. Auf ähnliche Weise lassen sich Umkehrungen von **9** beweisen. Wir fassen die damit insgesamt sich ergebenden Sätze zusammen in der Gleichung der W-Schreibweise:

9a. $W(O, P) \cdot W(O.P, Q) = W(O, Q) \cdot W(O.Q, P)$

9a nennen wir auch *Produktregel*.

10. $(O \underset{p}{\rightarrow} P).(O.P \underset{u}{\rightarrow} Q).(O.\bar{P} \underset{u'}{\rightarrow} Q) > (O \underset{q}{\rightarrow} Q).[q = p \cdot u + (1 - p) \cdot u']$

10a. $W(O, Q) = W(O, P) \cdot W(O.P, Q) + W(O, \bar{P}) \cdot W(O.\bar{P}, Q)$

Beweis durch Bildung von $(O \underset{r}{\rightarrow} P.Q)$ und $(O \underset{s}{\rightarrow} \bar{P}.Q)$ nach IV, 1, Anwendung von III, 1 hierauf und Benutzung von $P.Q \vee \bar{P}.Q = [P \vee \bar{P}].Q = Q$. Wir nennen **10** auch *Ablösungssatz*, da hier ein Mittel angegeben wird, die Wahrscheinlichkeit von Q in bezug auf O von P abzulösen.

11. $$(O \underset{p}{\rightarrow} P).(O \underset{q}{\rightarrow} Q).(O.P \underset{u}{\rightarrow} Q) > (O.\bar{P} \underset{u'}{\rightarrow} Q).\left(u' = \frac{q - p \cdot u}{1 - p}\right)$$

Zum Beweis bildet man $(O \underset{q}{\rightarrow} Q.[P \vee \bar{P}])$, umgeformt $(O \underset{q}{\rightarrow} Q.P \vee Q.\bar{P})$. Dies zerlegt man nach Satz **1**, unter Berücksichtigung von IV, 1, und bildet $(O \underset{s}{\rightarrow} \bar{P}.Q).(s = q - p \cdot u)$. Dies zerlegt man nach Satz **2**, wobei Satz **11** entsteht. Dieser Satz ist eine Umkehrung von Satz **10**; auch hier bewirken also die Umkehraxiome III, 2 und IV, 2 bis 3, daß man **10a** als Gleichung auffassen kann und nach einer beliebigen der auftretenden Größen $W(\ )$ auflösen darf.

Satz **11** führt auf zwei bemerkenswerte Spezialfälle, indem man einmal $p = q$ und das andre Mal $u = u'$ setzt:

12. $$(O \underset{p}{\rightarrow} P).(O \underset{p}{\rightarrow} Q).(O.P \underset{u}{\rightarrow} Q) > (O.\bar{P} \underset{u'}{\rightarrow} Q).\left(u' = \frac{p \cdot (1 - u)}{1 - p}\right)$$

In W-Schreibweise: Ist $W(O, P) = W(O, Q)$, so gilt:

12a. $$\frac{W(O.\bar{P}, Q)}{W(O.P, \bar{Q})} = \frac{W(O, P)}{W(O, \bar{P})} = \frac{W(O, Q)}{W(O, \bar{Q})}$$

13. $$(O \underset{p}{\rightarrow} P).(O \underset{q}{\rightarrow} Q).(O.P \underset{u}{\rightarrow} Q).(O.\bar{P} \underset{u}{\rightarrow} Q) > (u = q)$$

In W-Schreibweise: Ist $W(O.P, Q) = W(O.\bar{P}, Q)$, so gilt:

13a. $$W(O.P, Q) = W(O.\bar{P}, Q) = W(O, Q)$$

Aus Satz **11** folgt nun noch eine einschränkende Bedingung für die Wahl der vorgebbaren Wahrscheinlichkeitsgrade. Alle vier Größen p, q, u, u' müssen ja Satz 5 befriedigen. Gibt man nun drei von diesen Größen, etwa p, q, u, willkürlich vor, so kann der Fall eintreten, daß $u' < 0$ oder $u' > 1$ wird, obgleich jede der gegebenen drei Größen p, q, u für sich zwischen 0 und 1 (einschließlich der Grenzen) liegt. Wir erhalten deshalb den Satz:

14. $$(O \underset{p}{\rightarrow} P).(O \underset{q}{\rightarrow} Q).(O.P \underset{u}{\rightarrow} Q) > \left(0 \leqq \frac{q - p \cdot u}{1 - p} \leqq 1\right)$$

Dies läßt sich in zwei Ungleichungen zerspalten, die wir gleich in der W-Schreibweise wiedergeben:

14a. $$W(O.P, Q) \leqq \frac{W(O, Q)}{W(O, P)}$$

14b. $$\frac{W(O, Q) - W(O, P) \cdot W(O.P, Q)}{1 - W(O, P)} \leqq 1$$

Aus diesen Ungleichungen lassen sich entsprechende Ungleichungen für eine andre Wahl gegebener Größen ableiten.

15. $(O \underset{p}{\rightarrow} P).(O.P \underset{u}{\rightarrow} Q).(O.\bar{P} \underset{u'}{\rightarrow} Q) > (O.Q \underset{v}{\rightarrow} P).\left(v = \frac{p \cdot u}{p \cdot u + (1-p) \cdot u'}\right)$

15a. $$W(O.Q, P) = \frac{W(O, P) \cdot W(O.P, Q)}{W(O, P) \cdot W(O.P, Q) + W(O, \bar{P}) \cdot W(O.\bar{P}, Q)}$$

Zum Beweis beachten wir, daß die linke Seite von **15** identisch ist mit der linken Seite von **10**. Wir schließen nun zunächst mit **10** auf $(O \underset{q}{\rightarrow} Q)$. Außerdem folgt aus der linken Seite $(O \underset{w}{\rightarrow} P.Q).(w = p \cdot u)$. Dann benutzen wir Satz **2**, nachdem wir dort P mit Q vertauscht haben. — **15** ist die sogenannte *Bayessche Regel*; dies wird besonders in der Schreibweise **15a** erkenntlich. Man beachte, wie diese Formel aus der doppelten Zerlegbarkeit der Kombinationswahrscheinlichkeit hervorgeht, die wir in der Produktregel **9a** ausgesprochen haben; hinzu tritt noch der Ablösungssatz **10**.

16. $$(O \underset{p}{\rightarrow} P).(O \underset{q}{\rightarrow} Q).(O.P \underset{u}{\rightarrow} R).(O.Q \underset{v}{\rightarrow} R).(O.P > \bar{Q}) > (O.[P \vee Q] \underset{w}{\rightarrow} R).\left(w = \frac{p \cdot u + q \cdot v}{p + q}\right)$$

16a. $$W(O.[P \vee Q], R) = \frac{W(O, P) \cdot W(O.P, R) + W(O, Q) \cdot W(O.Q, R)}{W(O, P) + W(O, Q)}$$

Bei der Schreibweise **16a** ist die Ausschlußbedingung $(O.P > \bar{Q})$ inhaltlich hinzuzudenken. Zum Beweis schließen wir zunächst aus der linken Seite von **16** mit IV,1 bzw. III,1

$$(O \underset{r}{\rightarrow} P.R).(r = p \cdot u) \quad (O \underset{s}{\rightarrow} Q.R).(s = q \cdot v) \quad (O \underset{t}{\rightarrow} P \vee Q).(t = p + q)$$

Aus den beiden ersten dieser Formeln schließen wir mit III,1 auf $(O \underset{x}{\rightarrow} P.R \vee Q.R)$, was wir umformen in $(O \underset{x}{\rightarrow} [P \vee Q].R)$. Hieraus und aus der dritten Formel schließen wir mit Satz **2** auf $(O.[P \vee Q] \underset{w}{\rightarrow} R)$. –

Wir nennen **16** auch „Regel des vorderen oder“, weil diese Formel ein „oder“ im Vorderglied enthält. Es ist bemerkenswert, daß sich keine Formel aufstellen läßt, in welcher das Vorderglied aus einer Disjunktion *allein* besteht, ohne danebenstehenden Faktor O; d. h. genauer: eine derartige Formel läßt sich nicht aus Formeln ableiten, die diesen Charakter noch nicht besitzen. Schreibt man eine Implikation hin, deren rechte Seite eine Wahrscheinlichkeitsimplikation mit einer Disjunktion als alleinigem Vorderglied ist, deren linke Seite aber derartige Formeln nicht enthält, so ist das Axiomensystem in bezug auf den ganzen Ausdruck entscheidungsindefinit. Aus der Häufigkeitsdeutung wird dies verständlich, weil $P \vee Q$

allein nichts darüber besagt, in welchem Häufigkeitsverhältnis P und Q auftreten; eine Wahrscheinlichkeitsimplikation mit einer Disjunktion als alleinigem Vorderglied besitzt deshalb keinen bestimmten Wahrscheinlichkeitsgrad. Die Wahrscheinlichkeitsrechnung üblicher Darstellung spricht deshalb in derartigen Formeln wie **16** von einer „apriori-Wahrscheinlichkeit" von P und Q; diese apriori-Wahrscheinlichkeiten, die übrigens auch bei der Bayesschen Regel auftreten, rühren daher, daß man fälschlich die Wahrscheinlichkeit als Funktion *einer* Klassen-Variablen, anstatt als Funktion *zweier* Klassen-Variablen betrachtet hat. In unserer Darstellung tritt an Stelle dessen eine Wahrscheinlichkeit von P und Q in bezug auf O, und auf der rechten Seite muß O als Faktor neben die Disjunktion $P \vee Q$ treten.

Aus **16** ergibt sich noch ein bemerkenswerter Spezialfall für $u = v$:

17. $(O \underset{u}{\rightarrow} P).(O \underset{u}{\rightarrow} Q).(O.P \underset{u}{\rightarrow} R).(O.Q \underset{u}{\rightarrow} R).(O.P \supset \bar{Q}) \supset (O.[P \vee Q] \underset{u}{\rightarrow} R)$

In W-Schreibweise: Bestehen die Wahrscheinlichkeiten $W(O, P)$ und $W(O, Q)$, und ist $W(O.P, R) = W(O.Q, R)$, so gilt für einander ausschließende Ereignisse P und Q:

17a. $$W(O.P, R) = W(O.Q, R) = W(O.[P \vee Q], R)$$

Man beachte auch hier, daß das „oder" im Vorderglied nur auftreten kann, weil P und Q eine Wahrscheinlichkeit in bezug auf O besitzen.

18. $$(O \supset P).(O \underset{q}{\rightarrow} Q) \supset (O.P \underset{q}{\rightarrow} Q)$$

Zum Beweis schließt man aus $(O \supset P)$ auf $(O.Q \supset P)$, hieraus mit II, 1 auf $(O.Q \underset{r}{\rightarrow} P).(r = 1)$, hieraus mit $(O \underset{q}{\rightarrow} Q)$ und IV, 1 auf $(O \underset{q}{\rightarrow} P.Q)$. Nun nimmt man $(O \supset P)$ hinzu, das man mit II, 1 in $(O \underset{p}{\rightarrow} P).(p = 1)$ verwandelt, und schließt mit Satz **2** auf $(O.P \underset{q}{\rightarrow} Q)$.

19. $$(O \supset P).(O \underset{q}{\rightarrow} Q) \supset (O \underset{q}{\rightarrow} [P \supset Q])$$

Zum Beweis ersetzen wir $P \supset Q$ durch $\bar{P} \vee Q$. Aus der linken Seite läßt sich $(O \underset{r}{\rightarrow} \bar{P}).(r = 0)$ ableiten, und da $(O.Q \supset \bar{P})$ gilt, so folgt die rechte Seite von **19** aus III, 1.

§ 5.

Geometrische Deutung des Axiomensystems.

Das gegebene Axiomensystem läßt sich als System impliziter Definitionen für das Zeichen $\underset{p}{\rightarrow}$ auffassen. Man darf jedoch nicht glauben, daß damit bereits derjenige Begriff erfaßt sei, den wir bei Anwendungen der

Wahrscheinlichkeitsrechnung unter Wahrscheinlichkeitsimplikation verstehen. Vielmehr erhält man auf diese Weise einen allgemeineren Begriff, der zwar gewisse strukturelle Züge mit dem Wahrscheinlichkeitsbegriff gemein hat, der aber ebensogut Objekten gänzlich anderer Art zugeordnet werden kann. Wir wollen dies erläutern, indem wir eine geometrische Deutung des Axiomensystems angeben.

Wir zeichnen (Fig. 1) ein Koordinatenkreuz und verstehen unter O, P, Q Punkte auf den Achsen x und y. O legen wir in den Mittelpunkt, P auf die Abszissenachse, Q auf die Ordinate, jedoch so, daß sie innerhalb der von O aus abgetragenen Strecke 1 liegen. Unter $W(O, P) = p$ verstehen wir die Strecke OP, unter $W(O, Q) = q$ die Strecke OQ. Die Angabe des Abhängigkeitszustandes der drei Größen O, P, Q bedarf nach den Axiomen IV noch einer dritten Angabe, als welche wir die Größe $W(O, P.Q) = w$ wählen wollen. Wir zeichnen dazu die Hyperbel $x \cdot y = w$, die wir Abhängigkeitshyperbel nennen. Die Größe $W(O.P, Q) = u$ erhalten wir dann, indem wir in P eine Senkrechte bis an die Hyperbel ziehen und von da eine Horizontale bis zum Schnitt der y-Achse führen, so daß u als die hierdurch bestimmte Strecke auf der y-Achse dargestellt wird; durch eine entsprechende Konstruktion, die von Q aus beginnt, erhalten wir $W(O.Q, P) = v$ als eine Strecke auf der x-Achse. Wegen einer bekannten Eigenschaft der Hyperbel ist dann

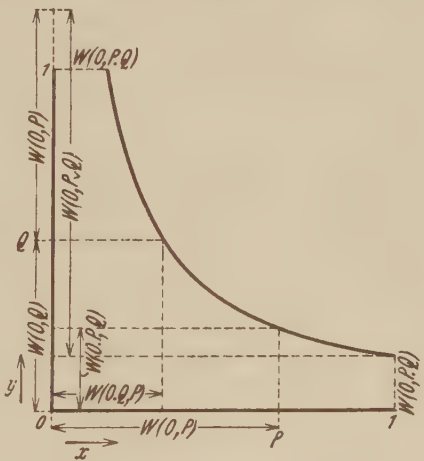


Fig. 1. Geometrische Deutung der Wahrscheinlichkeitsimplikation.

$$p \cdot u = q \cdot v = w$$

Die Größe w ist durch die Ordinate an der Stelle $x = 1$ bis zur Hyperbel, bzw. durch die Abszisse an der Stelle $y = 1$ bis zur Hyperbel dargestellt. Auch die Größe $W(O, P \vee Q) = r$, die durch Satz 8 bestimmt ist, findet in der Figur eine geometrische Darstellung als eine Strecke auf der y-Achse; ihr unterer Endpunkt ist durch eine Horizontale im Abstand w von der x-Achse bestimmt, ihr oberer Endpunkt dadurch, daß von Q die Strecke p nach oben abgetragen wird.

Man überzeugt sich leicht, daß durch die Konstruktion alle Axiome befriedigt werden. Eine Ausnahme bildet nur Axiom II, 1, insofern als der Ausdruck $O > P$, der ein logisches Zeichen enthält, keine Interpretation

findet; dagegen findet natürlich der Ausdruck $O \underset{1}{\rightarrow} P$ eine Interpretation, indem P in den Abstand 1 von O verlegt wird. In den Axiomen III, 1—2 ist der Ausdruck $(O.P \succ \bar{Q})$ für die geometrische Deutung durch $(O.P \underset{x}{\rightarrow} Q).(x = 0)$ zu ersetzen; man folgert mit Hilfe von Satz 8 leicht, daß diese Fassung der in III, 1—2 hingeschriebenen Fassung innerhalb des gesamten Axiomensystems äquivalent ist. — Bei dieser geometrischen Deutung sind also die Größen O, P, Q durch Punkte auf den Achsen innerhalb des Einheitsabstandes wiedergegeben, die verschiedenen Wahrscheinlichkeitsgrade durch Strecken auf den Achsen.

Die gegebene geometrische Deutung führt zu einer sehr anschaulichen Interpretation des Abhängigkeitsbegriffes. Der Abhängigkeitszustand der drei Punkte O, P, Q ist durch die Lage der Punkte im Koordinatensystem noch nicht bestimmt, sondern erst durch die hinzutretende Abhängigkeitshyperbel. In der Wahl der Hyperbel sind zwei Spezialfälle von Bedeutung. Der erste Spezialfall entsteht, wenn der Punkt $x = p$, $y = q$ auf der Abhängigkeitshyperbel liegt; dann ist $u = q$, $v = p$. In diesem Falle gilt also das spezielle Multiplikationstheorem, und die Ereignisse P und Q heißen „unabhängig voneinander in bezug auf O". Der andere Spezialfall entsteht, wenn die Abhängigkeitshyperbel ausartet und mit dem Koordinatenkreuz identisch wird. Dann ist $w = 0$, $u = 0$, $v = 0$; die beiden Ereignisse P und Q schließen sich aus in bezug auf O. Man erkennt, daß dann die Strecke $r = p + q$ wird und das spezielle Additionstheorem III, 1 in Kraft tritt.

Die Einführung der ausführlichen Schreibweise (1, § 2) bringt für die geometrische Deutung keine Schwierigkeiten mit sich; an Stelle der einen Figur tritt jetzt entsprechend dem laufenden Index eine Folge gleicher Figuren, gewissermaßen ein Film derartiger Figuren.

Die gegebene Belegung des Axiomensystems mit geometrischen Größen ist freilich eine etwas zu allgemeine Belegung, weil wir nur den ganzen Klammerausdrücken der Axiome, nicht den Einzelzeichen eine geometrische Deutung zugeordnet haben. Infolgedessen haben wir nicht nur die logischen Zeichen für „und" und „oder" ebenfalls geometrisch interpretiert, sondern auch keinen Gebrauch davon gemacht, daß z. B. der Punkt zwischen O und P in $(O.P \underset{p}{\rightarrow} Q)$ dasselbe Zeichen ist wie der Punkt zwischen P und Q in $(O \underset{w}{\rightarrow} P.Q)$. Das geometrische Modell realisiert deshalb einen etwas weiteren Begriff als den Begriff $\underset{p}{\rightarrow}$ des Axiomensystems. Andrerseits ist festzustellen, daß bei der Ableitung der Sätze das Axiomensystem ebenfalls nur in diesem allgemeinen Sinne benutzt wird da hier auch nur Um-

ormungen der ganzen Klammerausdrücke vorgenommen werden und z. B. von der erwähnten Identität des Punktzeichens im Vorder- und Hinterglied der Wahrscheinlichkeitsimplikation kein Gebrauch gemacht wird.

§ 6.

Arithmetische Deutung (Häufigkeitsdeutung) des Axiomensystems.

Wenn wir den Wahrscheinlichkeitsbegriff für gewöhnlich mit der Vorstellung einer Häufigkeit behaften, so muß dies vom Standpunkt der Axiomatik als eine inhaltliche Belegung des Axiomensystems aufgefaßt werden. Es entsteht dabei eine Deutung des Zeichens $\underset{p}{\rightarrow}$ durch die Häufigkeitseigenschaften von Folgen, also eine Abbildung des Axiomensystems auf arithmetische Verhältnisse, welche der in Fig. 1 gegebenen Abbildung auf geometrische Verhältnisse analog ist.

Um diese Deutung formulieren zu können, führen wir einen Hilfsbegriff ein. Wir verstehen unter $\overset{n}{\underset{1}{N}}_i(x_i \varepsilon A)$ die Anzahl derjenigen x_i zwischen $i = 1$ und $i = n$, welche der Klasse A angehören. Das Zeichen N ist entsprechend zu verstehen in Ausdrücken der Form $\overset{n}{\underset{1}{N}}_i(x_i \varepsilon A) . (y_i \varepsilon B)$, wo es die Anzahl der Paare x_i, y_i bezeichnet, die der hingeschriebenen Bedingung genügen; und ebenso in ähnlichen Ausdrücken. Unter $H_n(O, P)$ verstehen wir den Ausdruck

$$H_n(O, P) = \frac{\overset{n}{\underset{1}{N}}_i(x_i \varepsilon O) . (y_i \varepsilon P)}{\overset{n}{\underset{1}{N}}_i(x_i \varepsilon O)} \tag{1}$$

ler entsprechend auch für kompliziertere Fälle zu verstehen ist, derart daß die vor dem Komma in $H_n(O, P)$ stehenden Zeichen den Nenner des Bruches bestimmen, während im Zähler alle Zeichen auftreten. Wir geben nun die Häufigkeitsdeutung durch folgende Vorschrift:

Zuordnungsdefinition α. Unter $(O \underset{p}{\rightarrow} P)$ soll die Aussage verstanden werden, daß die Folgen x_i und y_i für die Häufigkeit $H_n(O, P)$ einen limes besitzen, der gleich p ist. In Zeichen

$$W(O, P) = \lim_{n=\infty} H_n(O, P) \tag{2}$$

Diese Häufigkeitsdeutung durch einen limes entspricht der durch Herrn v. Mises [12]) gegebenen Deutung des Wahrscheinlichkeitsbegriffs. Jedoch unterscheidet sich unsere Häufigkeitsdeutung dadurch, daß wir eine Regellosigkeitsforderung nicht aufstellen. Wir können darauf verzichten, weil wir von

[12]) R. v. Mises, Grundlagen der Wahrscheinlichkeitsrechnung, Math. Zeitschr. 5 (1919), S. 55.

vornherein das allgemeine Multiplikationstheorem zugrunde gelegt haben; den Begriff der „normalen Folge“ werden wir später als einen Spezialfall behandeln.

Infolgedessen können unsere Wahrscheinlichkeitsfolgen durch Vorschriften gegeben werden. Z. B. ist die Folge

$$P\,\bar{P}\,P\,\bar{P}\,P\,\bar{P}\ldots$$

eine Wahrscheinlichkeitsfolge in dem von uns definierten Sinne. Auch das von Poincaré angegebene Beispiel der letzten Ziffern einer Logarithmentafel[13]) ist für uns ein Beispiel einer Wahrscheinlichkeitsfolge. Wir müssen die Angebbarkeit der Folge durch eine Vorschrift hier verlangen, weil wir eine mathematische Häufigkeitsdeutung entwickeln, und eine unendliche mathematische Folge nicht anders als durch eine Vorschrift gegeben werden kann. Wir nennen den dabei entstehenden Wahrscheinlichkeitsbegriff, also den durch das Axiomensystem einschließlich der Zuordnungsdefinition α bestimmten Begriff, den *inhaltlichen mathematischen Wahrscheinlichkeitsbegriff*, im Gegensatz zu dem *formalen mathematischen Wahrscheinlichkeitsbegriff*, der durch das Axiomensystem allein definiert wird, also ohne eine inhaltliche Belegung. Daß auch die inhaltliche mathematische Wahrscheinlichkeitsrechnung das Wahrscheinlichkeitsproblem der Physik noch nicht vollständig erfaßt, werden wir später noch zeigen.

Für die inhaltliche mathematische Wahrscheinlichkeitsrechnung gewinnt nun das Axiomensystem eine sehr einfache Rechtfertigung. *Es stellt sich nämlich heraus, daß unter Zugrundelegung der Zuordnungsdefinition α die sämtlichen Axiome Tautologien werden.*

Für die Axiome I bis II ist dies ohne weiteres ersichtlich; es ist zu beachten, daß hier bei der arithmetischen Deutung auch das Axiom II, 1 tautologisch wird, das wir bei der geometrischen Deutung fortlassen mußten.

Für das Additionstheorem III ist der tautologische Charakter bei Benutzung der Häufigkeitsdeutung bekannt, wenn P und Q derselben Folge angehören; man erkennt jedoch leicht, daß unter Zugrundelegung unserer Definition (1) das gleiche gilt, wenn P und Q verschiedenen Folgen angehören. Es ist nämlich nach Definition (1)

$$(3)\qquad H_n(O, P \vee Q) = \frac{\overset{n}{\underset{1}{N_i}}(x_i \varepsilon O) \,.\, [(y_i \varepsilon P) \vee (z_i \varepsilon Q)]}{\overset{n}{\underset{1}{N_i}}(x_i \varepsilon O)}$$

Wenn $O \,.\, P \supset \bar{Q}$ gilt, ist dies gleich

$$\frac{\overset{n}{\underset{1}{N_i}}(x_i \varepsilon O) \,.\, (y_i \varepsilon P)}{\overset{n}{\underset{1}{N_i}} x_i \varepsilon O} + \frac{\overset{n}{\underset{1}{N_i}}(x_i \varepsilon O) \,.\, (z_i \varepsilon Q)}{\overset{n}{\underset{1}{N_i}}(x_i \varepsilon O)}$$

[13]) Poincaré, Calcul des Probabilités (Paris: Gauthier-Villars, 1912), S. 343.

nd es wird dann

3a) $$H_n(O, P \vee Q) = H_n(O, P) + H_n(O, Q)$$

Diese Gleichung bleibt beim Übergang zum limes erhalten, und es ist daher ür sich ausschließende Ereignisse

$$W(O, P \vee Q) = W(O, P) + W(O, Q)$$

Die Bedingung $O.P \supset \bar{Q}$ genügt also für die Addition der Wahrscheinlichkeiten. — Für das Umkehraxiom III, 2 läßt sich der Schluß in genau derselben Weise durchführen.

Der tautologische Charakter gilt aber ebenfalls für das Multiplikationstheorem IV, 1. Es ist nämlich nach Definition (1)

4) $$H_n(O, P.Q) = \frac{\overset{n}{\underset{1}{N_i}}(x_i \varepsilon O).(y_i \varepsilon P).(z_i \varepsilon Q)}{\overset{n}{\underset{1}{N_i}}(x_i \varepsilon O)}$$

$$= \frac{\overset{n}{\underset{1}{N_i}}(x_i \varepsilon O).(y_i \varepsilon P)}{\overset{n}{\underset{1}{N_i}}(x_i \varepsilon O)} \cdot \frac{\overset{n}{\underset{1}{N_i}}(x_i \varepsilon O).(y_i \varepsilon P).(z_i \varepsilon Q)}{\overset{n}{\underset{1}{N_i}}(x_i \varepsilon O).(y_i \varepsilon P)}$$

$$= H_n(O, P) \cdot H_n(O.P, Q)$$

Diese Gleichung bleibt für den Übergang zum limes gültig, wenn die einzelnen limites bestehen[13a]), und es ist daher mit (2) auch

5) $$W(O, P.Q) = W(O, P) \cdot W(O.P, Q)$$

Dies ist aber das allgemeine Multiplikationtheorem IV, 1. Hier liegt der Grund, warum die von uns gewählte Form IV, 1 des Multiplikationstheorems immer richtig ist. In dieser Form bedeutet nämlich das Multiplikationstheorem für die Häufigkeitsdeutung eine Tautologie. — Für die Umkehraxiome IV, 2—3 läßt sich der Schluß in genau derselben Weise durchführen; hierbei ist auf die früher allgemein genannte einschränkende Bedingung zu achten, daß die Wahrscheinlichkeit 0 für die Nennergrößen auszuschließen ist, da der Übergang von den Häufigkeiten zu den Wahrscheinlichkeiten sonst nicht möglich ist. Für den Fall, daß die Folge x_i überhaupt keine Glieder $x_i \varepsilon O$ enthält, werden die Häufigkeitsausdrücke $= \frac{0}{0}$; es steht uns frei, auch für diesen an sich unbestimmten Fall die Geltung der Axiome zu postulieren.

Die Frage nach der Herkunft der Axiome bildet also für die inhaltliche mathematische Wahrscheinlichkeitsrechnung kein Problem. Die Axiome

---

[13a]) Es genügt natürlich, dies für die rechte Seite zu fordern, da die Existenz des limes für die linke Seite dann beweisbar ist; eben darin besteht die Erfüllung des Axiom IV, 1. Das Entsprechende gilt für (3a).

formulieren tautologische Umformungen, die alle aus der Zuordnungsdefinition α folgen. *Die inhaltliche mathematische Wahrscheinlichkeitsrechnung ist nichts als ein Teilgebiet der Arithmetik, welches gewisse Umformungen von Folgen mit Limeseigenschaften in mathematisch strengen Methoden behandelt.*

Zugleich ist durch den Aufweis dieses arithmetischen Modells die Widerspruchslosigkeit der formalen mathematischen Wahrscheinlichkeitsrechnung, also des Axiomensystems von § 3, bewiesen.

II. Spezielle Fälle der mathematischen Wahrscheinlichkeitsrechnung.

§ 7.

Kritik des Regellosigkeitsprinzips.

Eine Reihe von Wahrscheinlichkeitsproblemen behandelt Folgen von spezielleren Eigenschaften, als sie mit der Zuordnungsdefinition α festgelegt werden. Es entsteht die Aufgabe, die Struktur derartiger Folgen durch zusätzliche Forderungen zu kennzeichnen.

Aus solchen Überlegungen heraus hat Herr v. Mises den interessanten Vorschlag gemacht, derartige Folgen durch ein Prinzip der Regellosigkeit zu kennzeichnen[14]). Dieses Prinzip besagt, auf Folgen von Alternativcharakter spezialisiert (vgl. § 1), daß jede beliebige Auswahl, welche von dem „Merkmal" P keinen Gebrauch macht, zu einer Teilfolge von gleichem limes der Häufigkeit führt. Diese Forderung können wir in unserem Kalkül auf folgende Weise wiedergeben. Eine Auswahl können wir durch eine Klasse X bezeichnen, welche durch eine Vorschrift $\varphi(x_i)$ gegeben ist, derart daß gewisse Elemente x_i, die O angehören, auch X angehören; z. B. kann X dadurch gegeben werden, daß jedes zweite Elemement der Klasse X angehört. Allgemeiner kann die Vorschrift φ auch auf gewisse Elemente z_i bezogen werden, die den x_i zugeordnet sind, derart daß X durch $\varphi(z_i)$ bestimmt ist. Die Forderung, daß X ohne Benutzung des Merkmals P definiert werden soll, können wir darstellen durch den Ausdruck

$$(1) \qquad \overline{X = F(P)}$$

den wir lesen können: „X ist keine Funktion von P." Dann nimmt die Forderung der Regellosigkeit die Form an

$$(2) \qquad (X)\,\overline{X = F(P)} \succ [W(O.X, P) = W(O, P)]$$

Es gelingt Herrn v. Mises aus diesem Prinzip eine Reihe wichtiger Eigenschaften für normale Folgen herzuleiten. Gegen das Prinzip der Regel-

[14]) R. v. Mises, Grundlagen der Wahrscheinlichkeitsrechnung, Math. Zeitschr. 5 (1919), S. 57.

osigkeit müssen jedoch einige tiefgehende Einwände erhoben werden, welche vir im folgenden zusammenfassen wollen.

Zunächst hat Herr v. Mises zeigen können, daß sein Prinzip zu der Forderung führt, eine Wahrscheinlichkeitsfolge könne nicht durch eine nathematische Vorschrift gegeben werden. Dieses Resultat stellt aber die Sinnhaftigkeit der benutzten Begriffsbildung in Frage. Über eine Folge, lie nicht durch eine mathematische Vorschrift gegeben werden kann, lassen ich auch keine mathematischen Aussagen machen. Die Zulassung solcher Aussagen führt in erhebliche Schwierigkeiten. Man kann etwa jeder Folge von Gliedern P und $\bar{P}$ einen unendlichen Dualbruch zuordnen, indem man für P, 0 für $\bar{P}$ setzt; dieser Dualbruch bestimmt eine reelle Zahl. Die v. Mises'schen Kollektivs sind damit der Klasse derjenigen reellen Zahlen ugeordnet, die sich durch keine Vorschrift geben lassen. Aber man wird vohl kaum bereit sein, eine reelle Zahl auf diese Weise zu charakterisieren; vom intuitionistischen wie vom formalistischen Standpunkt würde eine derartige Definition einer reellen Zahl gleichermaßen abzulehnen sein. Man könnte vielleicht sagen, daß die reelle Zahl durch keine mit endlichen Mitteln darstellbare Vorschrift angebbar sein soll; aber es ist bekannt, daß eine solche Begriffsbildung in Paradoxien führt.

Nun benutzt zwar v. Mises die genannte Eigenschaft seiner Kollektivs nicht zu ihrer Definition, sondern definiert mit Hilfe der Forderung, daß die auswählende Vorschrift von dem Merkmal P keinen Gebrauch machen soll; aber es scheint mir ebensowenig einen Weg zu geben, dieser Forderung einen mathematisch faßbaren Sinn zu erteilen. Gibt man nämlich eine Auswahlvorschrift an, so läßt sich durch nichts erkennen, ob sie nicht doch als Funktion des Merkmals hingestellt werden kann. Ist das Merkmal etwa durch eine recht komplizierte andersartige Vorschrift verteilt, so wird man zunächst die Auswahlvorschrift für unabhängig von der Merkmalverteilung halten, während sich für diesen Fall, wo eine Vorschrift für das Merkmal existiert, natürlich auch stets die Auswahlvorschrift als Funktion des Merkmals auffassen läßt. Ist das Merkmal aber, wie v. Mises fordert, in seiner Verteilung durch keine Vorschrift geregelt, so kann man die Unabhängigkeit der Auswahl von der Merkmalverteilung nur daran erkennen, *daß es uns nicht gelingt*, die Auswahl als Funktion des Merkmals darzustellen. Aber *dieser* Tatbestand ist natürlich auch schon für den erstgenannten Fall einer sehr komplizierten Merkmalverteilung erfüllt. Eine derartige Charakterisierung ist eben kein mathematisch benutzbares Verfahren, so wenig es etwa eine mathematische Charakterisierung ist, wenn man eine Zahl dadurch kennzeichnen wollte, daß man die Quadratwurzel aus ihr nicht durch bloßes Kopfrechnen ziehen kann. Aus diesem Grunde scheint mir die v. Mises'sche Charakterisierung der zulässigen Auswahl, und

damit auch unsere in (1) gegebene Formulierung dieses Verfahrens, keine mathematisch erlaubte Begriffsbildung zu sein. In der praktischen Anwendung wird man übrigens immer den umgekehrten Weg benutzen, und eine Auswahl dann unabhängig oder zulässig nennen, wenn sie den gleichen limes der Häufigkeit liefert; aber damit ist natürlich das Regellosigkeitsprinzip zu einer Leeraussage geworden, die von *jeder* Folge erfüllt wird.

Will man dem Prinzip der Regellosigkeit einen Sinn erteilen, so muß man es — und dies ist wohl auch die Auffassung von Herrn v. Mises — als eine nichtmathematische Aussage betrachten. Man kann nämlich behaupten, daß es uns Menschen nie gelingen wird, eine die Folge charakterisierende Vorschrift zu finden. Genauer gesagt: Konstruieren wir auf Grund einer gegebenen endlichen Zahl von Gliedern der Folge eine Vorschrift, so wird bei weiterer Verfolgung die Vorschrift sich nicht bewähren. Dies ist aber eine Aussage über die Beziehung von Folgen zu menschlichen Denkakten, wobei die Denkakte als Naturvorgänge anzusehen sind. Unter der Aussage (1) ist dann zu verstehen, daß die Klasse X nicht in *bewußter* Benutzung der Klasse P hergestellt wird, und (2) behauptet dann, daß die Klasse P von einer derartigen Klasse X unabhängig ist. Dies ist eine sinnvolle Aussage, vergleichbar etwa der Aussage, daß eine Würfelfolge von einer andern Würfelfolge unabhängig ist. Nur ist sie keine mathematische Aussage, denn die Beziehung mathematischer Größen zu menschlichen Denkakten geht in die Mathematik nicht ein. Die Mathematisierung des Regellosigkeitsgedankens ist in dem v. Misesschen Prinzip deshalb nicht geglückt.

Man könnte vermuten, daß trotzdem das Regellosigkeitsprinzip für die Anwendung der Wahrscheinlichkeitsrechnung auf physikalische Dinge brauchbar sei; (2) würde dann eine Aussage über die Unabhängigkeit der physikalisch hergestellten Folge y_i von andern physikalisch herstellbaren Folgen X bedeuten, wobei die Herstellung durch menschliche Denkakte einbegriffen ist. Es stellt sich jedoch heraus, daß das Prinzip (2) in diesem Falle zwar nicht *sinnlos*, wohl aber *falsch* ist. Es lassen sich nämlich immer Auswahlen angeben, welche (2) widersprechen. Um ein Beispiel zu geben: Würde man etwa bei einer Würfelfolge die Anfangslagen der Würfel, die Stärke des Stoßes, die Länge des Weges mit sehr genauen Instrumenten beobachten, so ließe sich, während der Würfel noch rollt, also P uns noch unbekannt ist, mit einiger Sicherheit voraussagen, ob P eintritt oder nicht. Wir stehen zwar nicht auf dem Standpunkt des Determinismus, daß sich dies bei hinreichender Verschärfung in der Beobachtung der Anfangsbedingungen mit Gewißheit, oder auch nur mit beliebiger Genauigkeit, vorausberechnen lassen müsse[15]); aber daß sich die Wahrscheinlichkeit der

[15]) Vgl. hierzu [5]).

Vorausberechnung wesentlich verbessern läßt, dürfte unbestreitbar sein. Denn es handelt sich ja hier noch um durchaus makroskopische Vorgänge. Und die aus der Quantenmechanik resultierende Einschränkung der Voraus-bestimmbarkeit spielt hier quantitativ noch keine nennenswerte Rolle. Die Hinzunahme der geschilderten Kenntnis über die Anfangsbedingungen des Wurfes läßt sich aber als Hinzutreten einer Klasse X zu O auffassen der-art, daß

$$\overline{X = F(P)} \,.\, [W(O \,.\, X, P) \neq W(O, P)] \tag{3}$$

gilt; dies widerspricht aber (2).

Man könnte versuchen, hiergegen einzuwenden, daß die Folge X unter bewußter Benutzung des Merkmals hergestellt sei. Aber dieser Einwand ist nicht stichhaltig. Zunächst haben wir schon darauf hingewiesen, daß bei der Herstellung von X das Merkmal P noch gar nicht bekannt ist, also auch nicht selbst benutzt werden kann. Wenn man sich aber darauf beruft, daß man auf Grund früherer Erfahrungen die Folge X als eine gewisse Funktion von P erkannt hat, so ist darauf zu erwidern, daß man diese Erkenntnis eben nur deshalb gewonnen hat, *weil* $W(O \,.\, X, P) \neq W(O, P)$ ist. Denn die physikalischen Experimente, welche zu der Festlegung von X führen, benutzen letzten Endes nichts als diesen Gedanken; man erkennt physikalisch die Abhängigkeit des P von den Anfangsbedingungen allein an der bestehenden Vorzugswahrscheinlichkeit. Auch hier gilt also wieder, daß wir die Auswahl nur deshalb abhängig von P nennen, weil sie eine ab-weichende relative Häufigkeit liefert. Diese Verteidigung des Regellosigkeits-prinzips verwandelt das Prinzip also wieder in eine Leeraussage, die von *jeder* Folge erfüllt wird.

Wir müssen deshalb das Prinzip der Regellosigkeit auch für eine an-gewandte Wahrscheinlichkeitsrechnung ablehnen, da es für physikalische Wahrscheinlichkeitsfolgen nicht zutrifft. Wenn Herr v. Mises sich bei seiner Begründung auf den Erfahrungssatz beruft, daß es bisher nicht geglückt sei, ein „Spielsystem" zu konstruieren, so wird dabei übersehen, daß dieses Resultat nur sehr beschränkt gültig ist. Mit den Mitteln zwar, die einem Spieler zur Verfügung stehen, ist bei den dort vorkommenden Folgen eine abweichende Auswahl nicht zu erzielen; mit den Mitteln aber, die dem Physiker zur Verfügung stehen, ist prinzipiell sehr wohl eine abweichende Auswahl erreichbar. (Vgl. hierzu § 8.)

Wie man es auch faßt, das Prinzip der Regellosigkeit ist stets ent-weder nicht sinnvoll, oder es ist eine Leeraussage, oder es ist falsch. Zur Begründung dieser Behauptung sei noch ein weiterer Versuch diskutiert, den man zur Deutung des Regellosigkeitsprinzips heranziehen könnte. Um die Formulierung zu vermeiden, daß das Merkmal keine Funktion von P

sein soll, könnte man etwa in folgender Weise den Begriff der zulässigen Auswahl zu definieren suchen: Man ordne jedem Glied y_i der Folge diejenige Folge zu, die entsteht, wenn y_i das gegenteilige Merkmal erhält, während alle andern Glieder ihr Merkmal behalten; dann soll, falls die Auswahl y_i trifft, in der zugeordneten Folge bei Benutzung derselben Auswahlvorschrift ebenfalls y_i getroffen werden. Dieser Gedanke trifft vielleicht am besten das, was man Unabhängigkeit der Auswahl vom Merkmal nennen will. Die Formulierung paßt auch für Beispiele; z. B. ist die Auswahl „Man wähle alle Glieder y_i, für die $y_i \varepsilon P$ gilt" danach keine zulässige Auswahl, während sich die Auswahl „Man wähle alle Glieder y_i, für welche $y_{i-1} \varepsilon P$ gilt" als eine zulässige Auswahl ergibt, wie es auch sein muß. In der Tat ist diese Definition der zulässigen Auswahl sinnvoll; jedoch läßt sich zeigen, daß es für jede Folge zulässige Auswahlen nach dieser Definition gibt, welche einen abweichenden Limes der Häufigkeit ergeben.

Zunächst erkennt man dies an dem geschilderten physikalischen Beispiel. Die Folge X der Anfangsbedingungen des Wurfes erfüllt das gegebene Kennzeichen der zulässigen Auswahl; dies rührt eben daher, daß die Anfangsbedingungen den Erfolg P nur mit Wahrscheinlichkeit bestimmen. Sodann läßt sich dieser Gedanke aber auch auf mathematische Folgen übertragen. Zwar gilt, daß unsere Definition der zulässigen Auswahl diesmal eine mathematisch faßbare Formulierung darstellt; aber das Regellosigkeitsprinzip ist dann eben auch für mathematische Folgen falsch. Dies erkennt man durch folgende Überlegung. Man stelle zunächst unter bewußter Benutzung des Merkmals eine einseitige Auswahl X her, für welche P eine einseitig abweichende Häufigkeit besitzt. Nun schreibe man die Auswahl X in Form einer unendlichen Tabelle hin, indem man neben jeden Index i a oder $\bar{a}$ schreibt, je nachdem ob i gewählt wird oder nicht (diese unendliche Tabelle ist genau so gut gegeben wie das Kollektiv selbst). Man verstehe unter der Auswahl X jetzt diese Tabelle. Dann ist diese Tabelle jetzt eine zulässige Auswahl, denn ihre Formulierung ist im Sinne der obigen Definition gegeben; die Auswahlvorschrift lautet jetzt etwa „Man wähle jedes y_i, für welches $a_i = a$ ist". In dieser Vorschrift wird von dem Merkmal kein Gebrauch gemacht, im Sinne der obigen Definition. Daß man vorher für die Gewinnung der Tabelle von dem Merkmal Gebrauch gemacht hatte, ist eine psychologische Angelegenheit, die in die mathematische Formulierung nicht eingeht. Und man erkennt auch den Grund für das Versagen der gegebenen Definition der zulässigen Auswahl: Diese Definition bestimmt nur die Art der *Formulierung* der Auswahl, aber sie legt die Auswahl selbst nicht fest, weil sich eine äquivalente Formulierung (durch die Tabelle) geben läßt, die die unzulässige Auswahl in eine zulässige Auswahl verwandelt.

Aus den angeführten Gründen müssen wir deshalb das Prinzip der gellosigkeit als unhaltbar bezeichnen. Der Versuch, die normalen Folgen rch Unabhängigkeit gegenüber Auswahlen eines sehr umfassenden Typs kennzeichnen, ist mißglückt; wir müssen uns deshalb darauf beschränken, ese Folgen durch eine Unabhängigkeit gegenüber Auswahlen eines sehr el engeren, dafür aber scharf definierten Typs zu charakterisieren[16]). Im genden wollen wir einen derartigen Weg benutzen.

§ 8.

Normale Folgen.

Wir benutzen zu unserer Charakterisierung den bisher in der Wahrheinlichkeitsrechnung üblichen Gedanken, die normale Folge durch Gelng des speziellen Multiplikationstheorems in der Aufeinanderfolge ihrer ieder zu kennzeichnen. Ehe wir diesen Gedanken durchführen, sind einige rbemerkungen vorauszuschicken.

Zunächst wollen wir unsere Überlegungen dadurch vereinfachen, daß wir e Folge x_i als aus lauter Gliedern $x_i \varepsilon O$ bestehend annehmen; auch soll ese Folge unendlich viele Glieder enthalten. Dadurch vereinfacht sich e zu gebende Begriffsbestimmung. Unsere Begriffsbildung bringt es nämh mit sich, daß wir zunächst nur ein normales Folgen*paar* definieren nnen. Indem wir nun festsetzen, daß die erste Folge dieses Paares den gegebenen Charakter haben soll, reduziert sich die weitere Begriffsstimmung des Paares auf Angaben über die Folge y_i. Wir werden deslb kurz von der normalen Folge y_i sprechen dürfen, indem wir dabei raussetzen, daß diese Folge als eingebettet in ein normales Folgenpaar betrachten ist.

Sodann haben wir eine Bemerkung über die Zählung zu machen. Es ot zwei Formen der Zählung für die Wahrscheinlichkeit von Kombinanen aufeinanderfolgender Glieder. Bei der *Zählung erster Art* bleibt der ufende Index für die Kombinationen derselbe wie in der Originalfolge; an zählt dann einzelne Glieder evtl. mehrfach, wenn sie entsprechend ftreten. So wird eine Aufeinanderfolge PPP zwei Kombinationen PP

[16]) Ein solcher Gedanke wird auch von K. Dörge (vgl. [2])) benutzt, dem es diesem Wege ebenfalls gelingt, das Regellosigkeitsprinzip zu eliminieren. Die von n begründete Wahrscheinlichkeitsrechnung stellt deshalb eine logisch korrekte ufigkeitstheorie dar, die dadurch, daß sie den heranzuziehenden Auswahltyp als riable behandelt, eine besonders elegante Form annimmt. Gegenüber der vorstehend wickelten Wahrscheinlichkeitsrechnung ist jedoch diese Theorie spezieller, weil sie h auf die Wahrscheinlichkeitsrechnung des speziellen Multiplikationstheorems beränkt; andererseits dürfte ihre Begriffsbildung Schwierigkeiten mit sich bringen, nn es sich um den Übergang zu physikalischen Anwendungen handelt.

liefern. In der ausführlichen Schreibweise (1, § 2) haben wir dies zu schreiben

$$(1) \qquad (i)\,([x_i \varepsilon O].[x_{i+1} \varepsilon O] \underset{w}{\rightarrow} [y_i \varepsilon P].[y_{i+1} \varepsilon P])$$

Wir wollen hierfür die Abkürzung

$$(2) \qquad (O.O^1 \underset{w}{\rightarrow} P.P^1)$$

einführen, indem wir die „Phase" dem Klassenbuchstaben als oberen Index hinzufügen. Dabei wollen wir die Regel einführen, daß vorn stets mindestens so viele Zeichen O, und in der entsprechenden Phase, stehen müssen, wie hinten Zeichen P stehen; weitere Zeichen O in anderen Phasen kann man wegen Satz 15 stets hinzufügen, da $(O > O')$ gilt. Zur Verdeutlichung der Schreibweise sei noch bemerkt, daß z. B.

$$(3) \qquad (O^r.P \underset{u}{\rightarrow} P^r)$$

identisch ist mit

$$(3\,a) \qquad (i)\,([x_{i+r} \varepsilon O].[y_i \varepsilon P] \underset{u}{\rightarrow} [y_{i+r} \varepsilon P])$$

Die neue Begriffsbildung der Phase bringt auch eine erweiterte Bedeutung unseres Axiomensystems mit sich. Wir setzen nämlich fest, daß die Zeichen O^r, P^r innerhalb der Axiome ebenso gebraucht und substituiert werden dürfen wie die Zeichen O, P, d. h. anders ausgedrückt, daß die Axiome in ausführlicher Schreibweise auch für Ausdrücke der Form (3a) gelten sollen.

Es läßt sich nun leicht zeigen, daß die Geltung des speziellen Multiplikationstheorems bei dieser Art der Zählung äquivalent ist mit der Forderung, daß die Auswahl durch eine beliebig vorgegebene Kombination von r Vorgängern in den Gliedern $r+1$ eine Teilfolge derart auswählt, daß in dieser Teilfolge für P derselbe limes der Häufigkeit besteht wie in der Hauptfolge. Wir wollen eine Folge mit der letzteren Eigenschaft als *nachwirkungsfrei* bezeichnen. Diese Eigenschaft können wir formulieren durch die Forderung

$$(4) \qquad (r)\,W(O^r.P^* \ldots P^{*r-1}, P^r) = W(O, P)$$

In dieser Schreibweise soll P^* bedeuten, daß man hierfür beliebig P oder $\bar{P}$ schreiben darf[17]). Hieraus folgt zunächst, unter Hinzunahme der eingangs erwähnten Festsetzung über die Folge x_i, daß auch die Folge y_i unendlich viele Glieder $y_i \varepsilon P$ enthalten muß, denn würde es nur eine endliche Zahl n von Gliedern $y_i \varepsilon P$ geben, so würde für $r > n$ eine Kombination

[17]) P^* ist also ein Kurzzeichen zur Vermeidung einer Konjunktion von Einzelausdrücken. Zum Verständnis sei noch bemerkt: Selbstverständlich darf man in allen Formeln P durch $\bar{P}$ ersetzen. Ist aber ein Phasenzeichen da, so darf man im allgemeinen dies nur tun, wenn man zugleich für alle P mit verschiedenen Phasenzeichen ebenfalls $\bar{P}$ einsetzt. Der Stern bedeutet, daß diese Bindung aufgehoben wird

von r Vorgängern P auf der linken Seite von (4) den Wert $\frac{0}{0}$ liefern, also keinen limes mehr bestimmen, so daß (4) ungültig würde. Infolgedessen gilt

$$W(O^r, P^r) = W(O, P) \tag{5}$$

da es für den limes nichts ausmacht, wenn endlich viele Glieder vorn weggelassen werden.

Aus (4) schließt man leicht auf die Geltung des speziellen Multiplikationstheorems für Kombinationen in der Zählung erster Art. Es ist nämlich nach Axiom IV, 1

$$\begin{aligned} &W(O\ldots O^r, P^*\ldots P^{*r}) \\ &= W(O\ldots O^r, P^*\ldots P^{*r-1})\cdot W(O\ldots O^r.P^*\ldots P^{*r-1}, P^{*r}) \\ &= W(O\ldots O^r, P^*\ldots P^{*r-1})\cdot W(O, P^*) \end{aligned} \tag{6}$$

Hieraus ergibt sich durch vollständige Induktion

$$W(O\ldots O^r, P^*\ldots P^{*r}) = W(O, P^*)\cdot W(O^1, P^{*1})\ldots W(O^r, P^{*r}) \tag{7}$$

was wegen (5) mit dem speziellen Multiplikationstheorem identisch ist.

Wenn man nach der Zählung erster Art die entstehenden Kombinationen, als Einzelglieder aufgefaßt, hintereinander schreibt, so ergibt sich, daß die dabei entstehende Folge nicht nachwirkungsfrei ist. So muß etwa das auf ein Glied $PPPP$ folgende Glied mit PPP beginnen, also entweder $PPPP$ oder $PPP\bar{P}$ lauten. Um dies zu vermeiden, benutzt man vielfach ein anderes Zählverfahren, indem man die Folge in Abschnitte der Länge r zerlegt und diese Abschnitte für sich durchzählt. Dies nennen wir Zählung zweiter Art.

Die *Zählung zweiter Art* können wir auf folgende Weise charakterisieren. Man spaltet die Hauptfolge durch die Vorschrift

$$i = k + m\cdot r \qquad (1 \leqq k \leqq r) \qquad (m = 0, 1, 2, 3, \ldots) \tag{8}$$

in r Teilfolgen, in welche jeweils m als laufender Index eintritt. Wir nennen dies eine *reguläre Spaltung*. Die einzelne Teilfolge ist durch die Zahl k festgelegt. Wir schreiben hier den Index k als unteren Index in O_k und P_k hinzu, also z. B., da k zwischen 1 und r variiert,

$$W(O_1\ldots O_r, P_1^*\ldots P_r^*) \tag{9}$$

In der ausführlichen Schreibweise würde dies lauten[17a])

$$(m)([x_{1,m}\,\varepsilon\, O]\,.\,[x_{2,m}\,\varepsilon\, O]\ldots[x_{r,m}\,\varepsilon\, O] \underset{w}{\rightarrow} [y_{1,m}\,\varepsilon\, P^*]\,.\,[y_{2,m}\,\varepsilon\, P^*]\ldots[y_{r,m}\,\varepsilon\, P^*]) \tag{10}$$

[17a]) Wir legen damit eine Erweiterung unserer Regel für den Übergang zur ausführlichen Schreibweise fest: an Stelle des Index i in den Elementen x_i und y_i tritt das durch (8) bestimmte Zahlenpaar k, m, und jedem Klassenzeichen O_k bzw. P_k ist ein Element $x_{k,m}$ bzw. $y_{k,m}$ mit demselben vorderen Index k zuzuordnen, während nur der hintere Index unter das Allzeichen tritt.

Bei der Zählung zweiter Art ist die Geltung des speziellen Multiplikationstheorems äquivalent mit der Forderung, daß die durch reguläre Spaltung entstehenden Teilfolgen die gleiche Wahrscheinlichkeit wie die Hauptfolgen besitzen und alle voneinander unabhängig sind.

Um dies zu beweisen, müssen wir zunächst den Begriff der Unabhängigkeit von Folgen untereinander kennzeichnen. Wir definieren: zwei Folgen P und Q heißen unabhängig voneinander in bezug auf die Folge O, wenn

$$(11) \qquad (r)\, W(O^r . P^{*i} \dots P^{*k}, Q^{*r}) = W(O, Q^*) \qquad (0 \leqq i, k \leqq r)$$

Die Erweiterung gegenüber der Unabhängigkeit von Ereignissen besteht darin, daß hier eine (3, § 3) entsprechende Bedingung für beliebige Phasen verlangt wird[18]). Entsprechend definieren wir die Unabhängigkeit von drei Folgen gegeneinander durch

$$(12) \qquad (r)\, W(O^r . P^{*i} \dots P^{*k} . Q^{*l} \dots Q^{*n}, R^{*r}) = W(O, R^*) \qquad (0 \leqq i, k, l, n \leqq r)$$

außerdem verlangen wir, daß diese Formel auch für jede Vertauschung der P, Q, R gelten soll. Man beachte, daß mit der Unabhängigkeit von P und Q in bezug auf O und der Unabhängigkeit von Q und R in bezug auf O noch nicht gesagt ist, daß auch P und Q und R in bezug auf O unabhängig sind. Lediglich die aus (12) zu gewinnenden Vertauschungen in bezug auf P, Q, R sind ableitbar, wenn man die Unabhängigkeit von je zwei der Folgen nach (11) voraussetzt; man nimmt dann zu (12) eine sinngemäße Erweiterung des Satzes 9a, § 4 hinzu. — Entsprechend definieren wir die Unabhängigkeit von r Folgen durch

$$(13) \qquad (r)\, W(O_r^r . P_1^{*i} \dots P_1^{*k} \dots P_{r-1}^{*l} \dots P_{r-1}^{*n}, P_r^{*r}) = W(O_r, P_r^*) \qquad (0 \leqq i, k, l, n \leqq r)$$

welcher Bedingung die durch Vertauschung der $P_1 \dots P_r$ untereinander entstehenden Formeln noch hinzuzufügen sind.

Man beweist nun leicht die Geltung des speziellen Multiplikationstheorems. Es ist nämlich

$$W(O_1 \dots O_r, P_1^* \dots P_r^*) = W(O_1 \dots O_r, P_1^* \dots P_{r-1}^*) \cdot W(O_1 \dots O_r . P_1^* \dots P_{r-1}^*, P_r^*)$$

und durch vollständige Induktion schließt man wegen (13) leicht, daß

$$(14) \quad W(O_1 \dots O_r, P_1^* \dots P_r^*) = W(O_1, P_1^*) \cdot W(O_2, P_2^*) \dots W(O_r, P_r^*)$$

Dies ist aber das spezielle Multiplikationstheorem.

---

[18]) Man erkennt hieraus zugleich, daß die Unabhängigkeit von Ereignissen im allgemeinen nur für eine bestimmte Phase besteht, während sie für eine andere Phase nicht zu bestehen braucht.

Wir definieren nun: *Eine Folge heißt normal, wenn bei jeder regulären Spaltung unabhängige Teilfolgen gleicher Wahrscheinlichkeit entstehen.*

Wir wählen dies als Definition der normalen Folge, weil sich hieraus zeigen läßt, daß normale Folgen auch nachwirkungsfrei sind, während das umgekehrte nicht notwendig gilt.

Für den Beweis, daß eine normale Folge nachwirkungsfrei ist, genügt es zu zeigen, daß in einer normalen Folge bei Zählung erster Art die Kombinationen durch das spezielle Multiplikationstheorem bestimmt sind. Es sei nun eine Kombination

$$P^* P^{*1} \dots P^{*r} \tag{15}$$

vorgegeben. Wir spalten dann die Folge regulär in r Teilfolgen. Die Kombination (15) kann dann auf r verschiedene Weisen aus den Teilfolgen hervorgehen; wir schreiben diese Anordnungen in den horizontalen Reihen des folgenden Schemas hin (der obere Index bedeutet hier die Phase in der Zählung der betreffenden Teilfolge):

$$\begin{array}{l} P_1^* \; P_2^* \; \dots P_{r-1}^* \; P_r^* \\ P_2^* \; P_3^* \; \dots P_r^* \; P_1^{*1} \\ P_3^* \; P_4^* \; \dots P_1^{*1} \; P_2^{*1} \\ \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \\ P_r^* \; P_1^{*1} \dots P_{r-2}^{*1} \; P_{r-1}^{*1} \end{array} \tag{16}$$

Wegen (13) besitzen diese Kombinationen alle gleiche Wahrscheinlichkeit x, und diese berechnet sich nach dem speziellen Multiplikationstheorem, gezählt nach zweiter Art. x ist also gleich dem in (14) rechts hingeschriebenen Ausdruck. Geht man jetzt, indem man rückwärts die Teilfolgen wieder zur Hauptfolge zusammensetzt, zur Zählung erster Art über, so multipliziert sich für den Zähler des Wahrscheinlichkeitsausdrucks die Größe x mit r, weil die Anzahl der Kombinationen nach (16) r mal so groß ist wie die Anzahl einer einzigen dieser Kombinationen; andererseits ist bei der Zählung erster Art der Nenner des Wahrscheinlichkeitsausdrucks auch r mal so groß. Daher ergibt sich für die Wahrscheinlichkeit einer derartigen Kombination in der Zählung erster Art wieder der durch (14) gegebene Wert x, und mit (5) ist der Satz bewiesen.

Weiter sieht man leicht, daß die regulären Teilfolgen einer normalen Folge selbst normal sind. Hat man nämlich eine Teilfolge, die durch r-fache Spaltung entstanden ist, so kann man diese Teilfolge dadurch in k weitere Teilfolgen spalten, daß man die Hauptfolge in $r \cdot k$ Teilfolgen spaltet. Da diese letzteren Teilfolgen wegen der Beziehung auf die Hauptfolge die Wahrscheinlichkeiten der Hauptfolge besitzen und voneinander unabhängig sind, so ist auch die erstgenannte Teilfolge normal.

Man kann den Begriff der normalen Folge noch enger spezialisieren, indem man eine weitere Forderung hinzufügt. Wir wollen diese „normale Folge im engeren Sinne“ jetzt noch definieren; die bisher definierte normale Folge bezeichnen wir dann auch als normale Folge im weiteren Sinne.

Man denke sich aus unendlich vielen normalen Folgen gleicher Häufigkeit ein Gitter gebildet, indem man die Folgen alle untereinander schreibt; bezeichnen wir hier das allgemeine Glied mit y_{ik}, wobei es wieder offen gelassen sein soll, ob $y_{ik} \varepsilon P$ oder $y_{ik} \varepsilon \bar{P}$ gilt, so nimmt ein derartiges Gitter die Form an

$$\begin{array}{l} y_{11}\, y_{21}\, y_{31} \cdots y_{i1} \cdots \\ y_{12}\, y_{22}\, y_{32} \cdots y_{i2} \cdots \\ \cdots\cdots\cdots\cdots \\ y_{1k}\, y_{2k}\, y_{3k} \cdots y_{ik} \cdots \\ \cdots\cdots\cdots\cdots \end{array} \tag{17}$$

Man kann die Folgen hier auch noch einzeln verschieben, in horizontaler Richtung; dann läßt man entsprechend bei einzelnen Folgen vorn endlich viele Glieder weg, bis die vordere Staffel lückenlos ist. Dabei darf man in der vorderen Staffel auch eine bestimmte Ordnung nach P herstellen, z. B. alle ersten Glieder gleich P machen. Wir definieren nun:

Ein System von Folgen y_{ik} (i laufend) besteht aus *normalen Folgen im engeren Sinne,* wenn zunächst alle Folgen y_{ik} (i laufend) normal im weiteren Sinne sind und für ihre Häufigkeiten H_{ik} den gleichen limes p haben, und wenn außerdem bei Gitterbildung auch alle Folgen y_{ik} (k laufend) für ihre Häufigkeit G_{ik} den limes p haben, abgesehen von der ersten Staffel; und zwar soll dies für jede beliebige Verschiebung gelten.

Die zusätzliche Eigenschaft ist für die normale Folge im weiteren Sinne nicht notwendig erfüllt[19]). Zur Veranschaulichung kann man sich etwa denken, daß in einem Gitter, welches aus normalen Folgen im weiteren Sinne besteht, an bestimmten Stellen einer jeden Horizontalfolge (also für vorgegebene Werte des vorderen Index, z. B. immer wenn dieser eine Primzahl ist) mit einer von p abweichenden, aber längs jeder Vertikalreihe konstanten Wahrscheinlichkeit gespielt wird. Sind diese Stellen hinreichend dünn gestreut, so macht sich diese Störung im limes der Häufigkeit nicht bemerkbar; auch die gegenseitige Unabhängigkeit aller Horizontalreihen ist mit dieser Störung vereinbar. Dagegen würde die Störung bemerkbar werden dadurch, daß die betreffenden Vertikalreihen einen abweichenden limes der Häufigkeit ergeben. Durch die Gitterbildung erhalten wir deshalb die Mög-

[19]) Übrigens erfüllen auch die v. Mises'schen Kollektivs nicht notwendig diese Forderung; man müßte dazu dem Regellosigkeitsprinzip eine entsprechende Erweiterung geben.

lichkeit, der Forderung konstanter Wahrscheinlichkeit von Glied zu Glied einen statistisch faßbaren Sinn zu erteilen.

Nur eine weit weniger besagende Eigenschaft der Vertikalreihen ist beweisbar, wenn die Horizontalreihen normal im weiteren Sinne und voneinander unabhängig sind. Es läßt sich nämlich durch Anwendung des Bernoullischen Theorems auf die ersten n Glieder einer Vertikalreihe, und weiter Grenzübergang zu $n = \infty$, zeigen, daß die (horizontal gezählte) relative Häufigkeit derjenigen Vertikalreihen, deren (vertikal gezählte) relative Häufigkeit im limes innerhalb eines beliebig kleinen Intervalls δ an p liegt, den limes 1 hat. D. h. die meisten Vertikalreihen haben einen limes der Häufigkeit nahe bei p. Aber damit ist noch nicht gesagt, daß auch nur eine einzige Vertikalreihe genau den limes p für ihre Häufigkeit hat. Das System der normalen Folgen im weiteren Sinne y_{ik} könnte nämlich auch ein konvergentes Gitter im Sinne von § 9 bilden.

Übrigens beachte man, daß unsere Definition der normalen Folge im engeren Sinne diesen Begriff nur für die horizontalen Folgen festsetzt; es läßt sich aus unserer Definition nicht folgern, daß auch die Vertikalfolgen normal sind. Dies müßte besonders postuliert werden, indem man die Vertikalfolgen als normal im weiteren Sinne annimmt.

Es ist zu beachten, daß die normale Folge im engeren Sinne nicht mehr durch explizite Definition, wie die normale Folge im weiteren Sinne, sondern nur durch implizite Definition bestimmt werden kann; d. h. es läßt sich nur die *Klasse* derartiger Folgen definieren, ob eine einzelne Folge dazu gehört, hängt davon ab, welche andern Folgen der Klasse angehören. Dies ist jedoch unter dem Gesichtspunkt der angewandten Wahrscheinlichkeitsrechnung kein Nachteil, da hier zumeist Vorschriften zur Definition von Wahrscheinlichkeitsfolgen benutzt werden, welche nicht die einzelne Folge, sondern eine Klasse von Folgen gleicher Wahrscheinlichkeit definieren (z. B. die Vorschrift des Würfelns)[20]. Die für diese Klasse von Folgen behauptete Wahrscheinlichkeitsimplikation ist dann die Wahrscheinlichkeitsimplikation im weiteren Sinne (vgl. § 2).

Mit dem Vorangehenden haben wir eine Definition der normalen Folge entwickelt, welche wesentlich enger ist als die mit dem Prinzip der Regellosigkeit gegebene und trotzdem für die in Anwendungen vorkommenden Fälle genügt. Gerade unter dem Gesichtspunkt der Anwendung trifft unsere Definition das Richtige, So wird der Gedanke der Unmöglichkeit des Spielsystems, den Herr v. Mises in dem Prinzip der Regellosigkeit formulieren wollte, durch den Begriff „nachwirkungsfrei“ ziemlich genau getroffen, denn was dem Spieler an Kenntnissen zur Verfügung steht, ist eine beobachtete

[20]) Dieser Gedanke ist von E. Tornier, Zeitschr. f. Physik 63 (1930), S. 697, in den Vordergrund gestellt worden; er kommt dadurch ebenfalls zu der Operation, die ich Gitterbildung genannt habe. Jedoch benutzt er keine Postulate, welche die Häufigkeit in den Vertikalfolgen auswerten. Ich selbst habe die Gitterbildung seit 1927 in meinen Vorlesungen benutzt.

Kombination von endlich vielen aufeinanderfolgenden Ereignissen P bzw. $\bar{P}$ und daß eine solche Kombination für das nachfolgende Ereignis keine Vorzugswahrscheinlichkeit ergibt, ist der Inhalt des Begriffs nachwirkungsfrei. Unsere Definition besitzt zudem den Vorzug, daß sie mit mathematischen Begriffen allein ausdrückbar ist; dies rührt daher, daß wir nur eine bestimmte mathematisch leicht formulierbare Auswahl benutzen, aber keinerlei Bezug auf menschliche Denkakte als Naturvorgänge nehmen. Ferner ist unsere Definition widerspruchsfrei, da das Bernoullische Theorem die Vereinbarkeit der Häufigkeitsdeutung für das Einzelereignis mit der Häufigkeitsdeutung für die Kombinationen beweist, wenn das spezielle Multiplikationstheorem zugrunde gelegt wird.

Trotzdem muß es als eine offene Frage bezeichnet werden, ob normale Folgen im Sinne unserer Definition, im engeren oder im weiteren Sinne, durch eine mathematische Vorschrift gegeben werden können. Sollte dies nicht der Fall sein, so würde dies bedeuten, daß in einer inhaltlichen mathematischen Wahrscheinlichkeitsrechnung der Begriff der normalen Folge nicht befriedigend definiert werden kann[21]).

Wir halten jedoch die ganze Frage der normalen Folge nicht für so wesentlich, zumal wenn man den Gesichtspunkt der Anwendung in den Vordergrund stellt. Ebensogut wie der Wahrscheinlichkeitsgrad einer Folge empirisch ermittelt werden muß, wird man auch für jede vorzugebende Auswahl den bei ihr entstehenden Wahrscheinlichkeitsgrad empirisch ermitteln müssen. Auf diese Weise wird man empirisch feststellen, in bezug auf welche Auswahl eine gegebene Wahrscheinlichkeitsfolge unabhängig ist, ebenso, ob eine Folge nachwirkungsfrei ist, auch in bezug auf welche Vorgänger, und ferner, ob zwei gegebene Folgen voneinander unabhängig sind. Die Folgen, mit denen wir es empirisch zu tun haben, besitzen zwar meistens eine Anzahl derartiger Auswahlen, stets aber gibt es auch Auswahlen von abhängigem Charakter. Und wenn auch der von uns definierte Begriff der normalen Folge bei Anwendungen benutzt wird, so hat man doch festzustellen, daß nichtnormale Folgen gerade in den Anwendungen mehr und mehr an Bedeutung gewinnen, da die Wahrscheinlichkeitsnachwirkung bei sehr vielen physikalischen Vorgängen eine Rolle spielt. Eine besondere Bedeutung besitzen hier Folgen, für welche der erste Vorgänger eine abweichende Auswahl liefert, während Hinzunahme weiterer

[21]) Man möchte jedoch vermuten, daß es möglich sein wird, eine normale Folge durch eine Vorschrift zu geben, da Borel (Traité du Calcul des Probabilités, Tome III Fasc. 1, Paris 1926, p. 1—10) bewiesen hat, daß die überwiegende Mehrheit aller Dezimalbrüche in ihren Ziffernfolgen normale Folgen (im weiteren Sinne) realisiert; er nennt diese Zahlen „nombres entièrement normaux“. Freilich gibt auch er kein Beispiel für einen solchen Dezimalbruch.

Vorgänger nichts mehr ändert. Solche Folgen besitzen durchaus Wahrscheinlichkeitscharakter im üblichen Sinne des Wortes; es läßt sich z. B. ein Bernoullisches Theorem für sie beweisen[23]). Sie sind im allgemeinen auch „zufallsartig"; diese Eigenschaft wird man den Folgen zweckmäßig immer dann zuschreiben, wenn keine der vorkommenden Auswahlwahrscheinlichkeiten gleich 0 oder 1 ist, denn dann läßt sich auf Grund einer vorliegenden Kombination das folgende Glied der Folge niemals mit Gewißheit voraussagen. Der Begriff der normalen Folge darf deshalb nur als Spezialfall innerhalb einer allgemeineren Wahrscheinlichkeitsrechnung gesehen werden, sonst würde man eine Reihe wichtiger Wahrscheinlichkeitsprobleme von vornherein aus der Wahrscheinlichkeitsrechnung ausscheiden.

§ 9.

Bernoullische Folgen.

Wir wollen noch einen andern Spezialfall definieren, der wesentlich umfassender ist als der der normalen Folgen und nahezu alle praktisch wichtigen Wahrscheinlichkeitsfolgen einschließen dürfte.

Die Folge x_i möge wieder aus lauter Elementen $x_i \varepsilon O$ bestehen, so daß wir die folgenden Bestimmungen nur für die Folge y_i zu treffen brauchen. Es sei nun nicht nur *eine* Folge y_i, sondern ein Gitter derartiger Folgen y_{ik} nach (17, § 8) vorhanden, derart daß für die Horizontalreihen (i laufend) alle Folgen die Wahrscheinlichkeit p realisieren. Im übrigen aber sei die Struktur der Folgen beliebig; sie brauchen also nicht normal zu sein. Wir definieren nun zunächst:

Ein System von Folgen y_{ik} bildet ein *konvergentes Gitter*, wenn alle Folgen y_{ik} (i laufend) für ihre Häufigkeit H_{ik} den limes p haben und die Folgen y_{ik} (k laufend) für ihre Häufigkeit Zahlen p_i zum limes haben, derart daß

$$\lim_{i=\infty} p_i = p \tag{1}$$

und zwar soll dies für jede beliebige Verschiebung gelten.

Diese Forderung besagt natürlich wesentlich weniger als die in § 8 benutzte, wo alle $p_i = p$ sind; sie ist auch von Folgen mit Wahrscheinlichkeitsnachwirkung erfüllbar und wird z. B. in der Gastheorie vorausgesetzt, wenn der Schluß von der Zeitgesamtheit auf die Raumgesamtheit vollzogen wird.

Wir führen nun einen weiteren Hilfsbegriff ein. Jedem Glied y_{ik} läßt sich die Häufigkeit H_{ik} der horizontalen Folge bis zur Stelle y_{ik} zuordnen.

---

[23]) Vgl. hierzu A. A. Markoff, Wahrscheinlichkeitsrechnung (Leipzig, Berlin 1912), S. 272–298; S. Bernstein, Math. Annalen 97 (1926), S. 1–59; G. Pólya, Sur quelques points de la theorie des probabilités II. Vorlesungen am Institut Henri Poincaré 1929.

Die H_{ik} lassen sich wieder einteilen, je nachdem $p - \delta \leqq H_{ik} \leqq p + \delta$ ist oder nicht; den ersteren Fall bezeichnen wir mit $H_{ik} \varepsilon K$, den andern mit $H_{ik} \varepsilon \bar{K}$. Wir erhalten dann ein Alternativ-Gitter der H_{ik}, welches dem y_{ik}-Gitter zugeordnet ist; wir bezeichnen es als das System der y_{ik} hinsichtlich K zugeordneten Häufigkeitsfolgen. In dem H_{ik}-Gitter können wir die Häufigkeit des Falles $H_{ik} \varepsilon K$ auszählen. Wir definieren nun:

Ein System von Folgen y_{ik} besteht aus *Bernoullischen Folgen*, wenn die den y_{ik} hinsichtlich K zugeordneten Häufigkeitsfolgen H_{ik} für jedes δ ein konvergentes Gitter bilden.

Die Bernoullischen Folgen sind, wie die normalen Folgen im engeren Sinne, nur durch implizite Definition charakterisierbar. Den Namen haben wir gewählt, weil sich für derartige Folgen zeigen läßt, daß sie ein Bernoullisches Theorem erfüllen. Wir beweisen dies folgendermaßen. In jeder horizontalen Reihe wird das H_{ik}-Gitter die Eigenschaft haben, daß von einer Stelle n ab alle weiteren Glieder nach K fallen; dies folgt aus der limes-Forderung für die horizontalen Reihen. (Es gibt zwar im allgemeinen kein n derart, daß dies für *alle* horizontalen Reihen zugleich gilt.) Das H_{ik}-Gitter nimmt also die folgende Gestalt an, wenn wir nur das Merkmal hinschreiben und diejenigen Glieder, in denen es unentschieden ist, ob K oder $\bar{K}$ vorliegt, mit K^* bezeichnen:

$$
\begin{array}{l}
K^* \; K^* \; K^* \; \ldots \; K \; K \; K \; K \; K \; \ldots \\
K^* \; K^* \; K^* \; \ldots \; K \; K \; K \; K \; \ldots \\
K^* \; K^* \; K^* \; \ldots \; K \; K \; \ldots \\
\ldots
\end{array} \tag{2}
$$

In den horizontalen Reihen ist hier der limes für die Häufigkeit von K gleich 1. Da H_{ik} in bezug auf K ein konvergentes Gitter bilden soll, muß dann die Häufigkeit in der senkrechten Staffel im limes durch Zahlen w_i gegeben sein, derart daß

$$\lim_{i=\infty} w_i = 1 \tag{3}$$

Nun bedeutet aber K den Fall, daß die Häufigkeit H_{ik} der horizontalen Folge an der i-ten Stelle innerhalb der Grenzen $\pm\delta$ bei p liegt; die hinsichtlich K geordnete senkrechte Staffel H_{ik} (k laufend) bedeutet dann die Aussonderung wiederholter Realisierungen von Kombinationen zu i Gliedern unter dem Gesichtspunkt, ob ihre Häufigkeit innerhalb $p \pm \delta$ liegt, und w_i bedeutet deshalb die Wahrscheinlichkeit, daß in einer Kombination von i Gliedern y_{ik} die Häufigkeit H_{ik} für den Fall $y_{ik} \varepsilon P$ innerhalb $p \pm \delta$ liegt; w_i ist also eine der bei Bernoulli auftretenden Wahrscheinlichkeiten, und (3) ist nichts anderes als die Behauptung des Bernoullischen Theorems, da derselbe Beweis sich für beliebiges δ durchführen läßt.

Es ist von Interesse, daß die limes-Eigenschaft der horizontalen Folgen für diesen Beweis noch nicht genügt, denn (3) läßt sich aus (2) nicht schließen, sondern bedeutet eine hinzutretende Annahme. Dagegen ist es nicht nötig, vorauszusetzen, daß auch die y_{ik} selbst schon ein konvergentes Gitter bilden.

Die Operation der Gitterbildung hat nun noch eine andere Bedeutung. Sie erlaubt nämlich, Wahrscheinlichkeitsfolgen zu definieren, in denen von Glied zu Glied mit einer andern Wahrscheinlichkeit gespielt wird. Für unsere ursprüngliche Definition (1, § 2), welche die Wahrscheinlichkeit p als Eigenschaft der ganzen Folge auffaßt, ist dies zunächst nicht möglich. Durch das Verfahren der Gitterbildung aber ist es möglich, die Wahrscheinlichkeit p_i auf die bisherige Weise als Eigenschaft der senkrechten Folge zu definieren, und damit wieder der horizontalen Folge eine von Glied zu Glied variable Wahrscheinlichkeit zuzuordnen. Für eine solche horizontale Folge gilt also

$$(4) \qquad (i)\,(x_{ik}\,\varepsilon\,O \underset{p_i}{\Rightarrow} y_{ik}\,\varepsilon\,P)$$

wo p_i eine Funktion von i ist. Hier hat also der Ausdruck

$$(5) \qquad x_{ik}\,\varepsilon\,O \underset{p_i}{\Rightarrow} y_{ik}\,\varepsilon\,P$$

scheinbar Selbständigkeit erreicht. Jedoch läßt sich die hier auftretende indivualistische Wahrscheinlichkeit auf eine Klassenwahrscheinlichkeit zurückführen, wenn man bedenkt, daß (4) nur innerhalb des ganzen Ausdrucks

$$(6) \qquad (i)\,(k)\,(x_{ik}\,\varepsilon\,O \underset{p_i}{\Rightarrow} y_{ik}\,\varepsilon\,P)$$

einen Sinn hat und aus diesem durch Vertauschung der Allzeichen und Weglassen von (k) entstanden ist.

Beispiele derartiger Folgen sind Folgen, in denen mit verschiedenen Wahrscheinlichkeiten $p_1 \ldots p_n$, die sich in n-fachem Turnus wiederholen, gespielt wird (Poissonsches Theorem), oder auch Folgen, bei denen mit Wahrscheinlichkeiten p_i gespielt wird, die für $i \to \infty$ nach p konvergieren. In letzterem Falle bilden die y_{ik} hinsichtlich P ein konvergentes Gitter. Handelt es sich in den horizontalen Folgen um Folgen mit Wahrscheinlichkeitsnachwirkung, so läßt sich ihre Entstehung aus verschiedenen Einzelwahrscheinlichkeiten schon innerhalb der einzelnen Folge erkennen. Jedoch kann die horizontale Einzelfolge auch normal im weiteren Sinne sein; dann ist der Folge ihre Entstehung aus Einzelereignissen verschiedener Wahrscheinlichkeiten allein gar nicht anzumerken, sie hat für sich betrachtet alle Eigenschaften einer Folge, die mit der konstanten Wahrscheinlichkeit p gespielt wird. Dies ist z. B. bei Folgen möglich, wo die y_{ik} hinsichtlich P ein konvergentes Gitter bilden. In diesem Falle wird erst bei Gitterbildung der veränderliche Charakter der Wahrscheinlichkeit von Glied zu Glied merkbar.

III. Angewandte Wahrscheinlichkeitsrechnung.

§ 10.

Das Sinnproblem der Wahrscheinlichkeit.

Die bisher entwickelte Wahrscheinlichkeitsrechnung war rein mathematischer Natur. Wenn wir jetzt zur angewandten Wahrscheinlichkeitsrechnung übergehen, d. h. zur Anwendung der mathematischen Disziplin auf wirkliche Dinge, so müssen wir eine Zuordnungsdefinition vorausschicken, welche angibt, was wir für die Anwendung auf wirkliche Dinge unter der Wahrscheinlichkeitsimplikation verstehen wollen. Als solche Zuordnungsdefinition wählen wir wieder Zuordnungsdefinition α mit der Zusatzbemerkung, daß die darin genannten Folgen nicht mathematisch gegebene Folgen, sondern *naturgegebene* Folgen sein sollen, wie sie etwa bei der Wiederholung physikalischer Vorgänge oder bei der gleichzeitigen Realisation von Zuständen in einer großen Menge von Exemplaren auftreten. Wir legen damit für die angewandte Wahrscheinlichkeitsrechnung von vornherein die Häufigkeitsdeutung fest. Für andere Deutungen der Wahrscheinlichkeit, die schon der individuellen Wahrscheinlichkeitsaussage einen Sinn zuschreiben wollen, sehen wir keinen gangbaren Weg. Beschränkt man sich nämlich darauf, die Wahrscheinlichkeitsaussage rein extensional, d. h. durch Eintreten oder Nichteintreten von Ereignissen, zu charakterisieren, so fehlt für die individuelle Aussage jede Möglichkeit, dem Wahrscheinlichkeitsgrad p eine objektive Deutung zuzuordnen. Dies gelingt erst, wenn p durch die Häufigkeit des Eintretens oder Nichteintretens von Ereignissen charakterisiert wird.

Mit dieser Deutung treten nun aber zwei Probleme in den Vordergrund, die innerhalb der mathematischen Wahrscheinlichkeitsrechnung noch keine Rolle spielen: dies sind das Sinnproblem und das Geltungsproblem der Wahrscheinlichkeitsaussagen.

Innerhalb der mathematischen Wahrscheinlichkeitsrechnung ist das Sinnproblem durch die vorangehenden Untersuchungen abschließend beantwortet. Wir fanden, daß man die Wahrscheinlichkeitsimplikation entweder als durch das Axiomensystem implizit definierten Begriff oder als limes-Charakter von Folgen deuten kann. Beides sind gleichberechtigte Deutungen. Daß die erste Deutung von logisch einwandfreiem Charakter ist, erkennt man aus ihrer Analogie zur mathematischen Deutung der Geometrie. Aber auch die zweite Deutung ist von logisch einwandfreiem Charakter, weil wir die mathematischen Wahrscheinlichkeitsfolgen grundsätzlich als durch eine Vorschrift gegeben betrachten können. In den Axiomen haben wir zwar keinerlei Voraussetzung über die Gebungsvorschrift der Folgen machen müssen, so daß die Axiome für alle Folgen gelten; aber wir werden ein bestimmtes Exemplar einer Wahrscheinlichkeitsfolge nur dann als vorliegend, als gegeben

betrachten können, wenn es durch eine Vorschrift gegeben wird. Infolgedessen erhalten alle Konvergenzaussagen über die Folgen den üblichen mathematischen Sinn. Es kann also grundsätzlich für jedes vorgegebene δ dasjenige n ausgemacht werden, für welches die Konvergenz innerhalb δ erreicht ist, derart, daß von da ab die Abweichung der Häufigkeit vom limes die Grenze δ nicht mehr überschreitet.

Innerhalb der angewandten Wahrscheinlichkeitsrechnung aber besteht für das Sinnproblem der Häufigkeitsdeutung eine grundsätzliche Schwierigkeit; sie rührt daher, daß Wahrscheinlichkeitsfolgen, welche durch Naturvorgänge geliefert werden, grundsätzlich auf eine andere Weise gegeben werden als mathematische Folgen. Und zwar liegen die Verhältnisse folgendermaßen. In der Folge x_i werden diejenigen Elemente, die zu O gehören, durch eben diese Forderung ausgewählt; die Eigenschaften dieser Folge sind daher, soweit sie für die Wahrscheinlichkeit gebraucht werden, durch die definierende Vorschrift bestimmt, d. h. intensional gegeben. In der Folge y_i aber werden diejenigen Elemente, die zu P gehören, durch die Natur geliefert; aus der definierenden Vorschrift folgt daher nichts über die benötigten Eigenschaften dieser Folge, d. h. diese Eigenschaften sind extensional gegeben. Infolgedessen läßt sich der Konvergenzaussage nicht die für mathematische Folgen geforderte Bestimmtheit erteilen; das zu einem δ gehörige n der Konvergenzstelle läßt sich nicht berechnen. Ja, es läßt sich auch nicht durch Ausprobieren bestimmen. Zwar läßt sich entscheiden, ob an einer Stelle n die Abweichung innerhalb δ liegt; um aber zu entscheiden, ob die Abweichung von da ab innerhalb δ *bleiben* wird, müßten wir die ganze unendliche Folge kennen.

Man hilft sich hier gewöhnlich durch die Fiktion eines idealen Beobachters, der die ganze extensional gegebene unendliche Folge überschaut und auf diese Weise die Konvergenzstellen feststellt. Aber es ist klar, daß eine derartige Fiktion nur dann erlaubt ist, wenn sie eliminierbar ist. Bei einer durch Vorschrift gegebenen Folge z. B. kann man sich durch diese Fiktion die limes-Eigenschaften veranschaulichen; aber eben nur deshalb, weil man der limes-Aussage auch ohne die Fiktion des idealen Beobachters einen Sinn erteilen kann. Für naturgegebene Folgen aber läßt sich die Fiktion des idealen Beobachters nicht eliminieren; und eben darum darf sie auch nicht benutzt werden. Dann aber scheint es unmöglich, der limes-Aussage einen Sinn zuzuschreiben; denn eine Aussage hat nur dann einen Sinn, wenn sie entscheidbar ist. Unter „entscheidbar“ hat man bisher immer die Forderung verstanden, daß ein beobachtbarer Tatbestand denkbar ist, welcher der Aussage das Prädikat wahr oder falsch erteilt. Da aber nach dem Vorangehenden für die in Wahrscheinlichkeitsaussagen geforderte Konvergenz dieses Urteil nie gefällt werden kann, weil stets nur eine endliche

Zahl von Fällen beobachtet werden kann, so scheint die limes-Aussage für naturgegebene Folgen sinnleer zu sein.

Diese Schwierigkeiten haben schon seit längerer Zeit die Philosophen exakter Richtung beschäftigt. Da ich die Antwort, die mir in dieser Lage die einzig mögliche zu sein scheint, an anderer Stelle[23]) ausführlich begründet habe, sei sie hier nur kurz referiert. Ich halte eine Auflösung dieses Problems innerhalb der strengen Logik für unmöglich, und bin deshalb zur Aufstellung einer allgemeineren Wahrscheinlichkeitslogik übergegangen. In der Wahrscheinlichkeitslogik wird der Grundsatz, daß jede Aussage wahr oder falsch sein muß, aufgegeben; an Stelle dessen wird es als hinreichend betrachtet, daß eine Aussage wahrscheinlich ist. Die zwei Wahrheitswerte der strengen Logik werden hier durch eine stetige Wahrheitsskala ersetzt, eben die Wahrscheinlichkeit, derart, daß Wahrheit und Falschheit nur die Grenzfälle der Wahrscheinlichkeitsskala bedeuten. Die strenge Logik erscheint danach, historisch betrachtet, als eine Idealisierung, zu der man durch den hohen Wahrscheinlichkeitsgrad der meisten Aussagen des täglichen Lebens gekommen ist, indem man eine hohe Wahrscheinlichkeit durch die Gewißheit ersetzte.

Innerhalb der Wahrscheinlichkeitslogik werden die Wahrscheinlichkeitsaussagen entscheidbar. Zwar sind sie auch hier nicht als wahr entscheidbar, aber sie sind *als wahrscheinlich entscheidbar*; das ist für die Wahrscheinlichkeitslogik sinnvoll und hinreichend. Zur Rechtfertigung unserer Auffassung verweisen wir darauf, daß dieser Gedanke dem tatsächlich beobachteten Verhalten im täglichen Leben wie in der Wissenschaft entspricht. Behauptet man etwa, daß die Wahrscheinlichkeit einer Würfelseite $^1/_6$ beträgt, und beobachtet man, daß unter 600 Würfen die Seite 1 in 96 Fällen vorgekommen ist, so wird man dies als eine induktive Bestätigung der Aussage betrachten; die Aussage wird durch den Beobachtungsbefund zwar nicht als wahr, aber doch als sehr wahrscheinlich entschieden. Wäre andererseits unter den 600 Fällen die Seite 1 nur viermal vorgekommen, so wäre die Aussage zwar nicht als falsch, aber doch als sehr unwahrscheinlich entschieden. Der Übergang von der Wahrheitslogik zur Wahrscheinlichkeitslogik macht also die limes-Aussage zu einer entscheidbaren und darum sinnvollen Aussage; freilich tritt damit der Wahrscheinlichkeitsbegriff aus dem Rahmen des Mathematischen heraus, denn Mathematik kennt nur Wahrheitslogik.

Diese Auflösung des Sinnproblems setzt jedoch noch die Anerkennung eines Axioms voraus, das wir Induktionsaxiom nennen. Da dieses Axiom schon in den Kreis des Geltungsproblems gehört, sei es in den folgenden Paragraphen verschoben.

---

[23]) H. Reichenbach, Kausalität und Wahrscheinlichkeit. Erkenntnis I, 1930, S. 168.

§ 11.

Das Geltungsproblem der Wahrscheinlichkeit.

Für die mathematische Wahrscheinlichkeitsrechnung enthält ebenso wie das Sinnproblem auch das Geltungsproblem keine Schwierigkeit. Die Axiome bedeuten hier lediglich willkürliche Definitionen; und zwar gilt dies sowohl für die formale wie für die inhaltliche mathematische Wahrscheinlichkeitsrechnung, da auch die in Zuordnungsdefinition α formulierte statistische Deutung eine willkürliche Festsetzung darstellt. Für die angewandte Wahrscheinlichkeitsrechnung dagegen entsteht die Frage nach der Geltung der Axiome.

Hier wird nun das Ergebnis des § 6 wichtig, da wir in der angewandten Wahrscheinlichkeitsrechnung stets die mit Zuordnungsdefinition α gegebene statistische Deutung der Wahrscheinlichkeit zugrunde legen. Wir hatten dort gefunden, daß mit der statistischen Deutung sämtliche Axiome I bis IV der Wahrscheinlichkeitsrechnung tautologisch werden. Für diese Axiome besteht deshalb kein Geltungsproblem anderer Art als das der Logik überhaupt, und wir brauchen sie deshalb in unsere erkenntnistheoretische Untersuchung nicht weiter einzubeziehen. Dagegen entsteht jetzt die Frage, was uns berechtigt, die in Zuordnungsdefinition α ausgesprochene statistische Deutung der Wahrscheinlichkeit von der Wirklichkeit zu behaupten. Diese Frage bedarf jedoch noch einer Richtigstellung. Wir behaupten zunächst mit Zuordnungsdefinition α nur: wenn eine Folge von limes-Charakter in der Wirklichkeit vorliegt, so wollen wir sie eine Wahrscheinlichkeitsfolge nennen. Dies ist nichts als eine Definition des Begriffs wahrscheinlich. Ein erkenntnistheoretisches Problem entsteht erst, wenn man von gewissen physikalisch herstellbaren Folgen behauptet, daß sie limes-Charakter haben. Aber gerade dafür müssen wir eine axiomatische Forderung formulieren, denn wir fanden in § 10, daß für eine physikalisch herstellbare Folge, da immer nur endlich viele Glieder realisiert werden können, der limes-Charakter nicht direkt festgestellt werden kann; die Behauptung, daß eine Naturfolge limes-Charakter besitzt, ist also keine Feststellung, sondern eine über das Feststellbare hinausgehende axiomatische Forderung.

Nach derartigen Anwendungsaxiomen der Wahrscheinlichkeit hat man in der Wahrscheinlichkeitsrechnung schon immer gesucht. Das Prinzip vom mangelnden Grunde z. B. gehört hierher; es knüpft an einen feststellbaren Tatbestand, das Fehlen von Gründen, die Behauptung der Gleichwahrscheinlichkeit, und für die statistische Deutung damit auch der Gleichhäufigkeit. Daß es sich in diesem Prinzip freilich um ein völlig unberechtigtes Postulat handelt, ist so einsichtig und auch so oft schon dargelegt worden, daß wir auf dieses Postulat hier nicht weiter einzugehen brauchen. Ein anderes

Prinzip der hier in Frage stehenden Art, also ein Anwendungsaxiom, ist das von mir früher aufgestellte Axiom der Wahrscheinlichkeitsfunktion[24]). Dieses Axiom erweist sich für zahlreiche Anwendungsfälle als gültig, insbesondere vermag es das Prinzip vom mangelnden Grunde vollständig zu ersetzen. Ich möchte jedoch heute dieses Prinzip durch ein noch allgemeineres Axiom ersetzen; das Prinzip der Wahrscheinlichkeitsfunktion ist, wie wir zeigen werden, als ein Spezialfall anzusehen, der für gewisse physikalische Tatbestände gilt, für diese aber auf Grund des neuen allgemeineren Axioms abgeleitet werden kann.

Um dieses allgemeinere Axiom formulieren zu können, führen wir den Begriff der *Häufigkeitsgröße* ein. Wir verstehen darunter eine Größe, die entweder eine relative Häufigkeit nach (1, § 6) ist, oder die Summe oder Differenz zweier derartiger Größen, oder der limes einer Folge derartiger Größen, oder auch eine Größe, die durch mehrfach wiederholte Anwendung der genannten Operationen in beliebiger Reihenfolge aus relativen Häufigkeiten entsteht. Das gesuchte Anwendungsaxiom lautet nun:

Axiom der Induktion. *Wenn eine physikalische Vorschrift eine beliebig verlängerbare Folge von Häufigkeitsgrößen h_i bestimmt, von denen bereits $h_1 \ldots h_n$ ermittelt vorliegen, und wenn s_1 und s_2 Zahlen sind derart, daß von einem Glied h_m $(1 \leqq m \leqq n)$ ab alle h_i $(m \leqq i \leqq n)$ zwischen den Schranken s_1 und s_2 liegen, so besteht eine Wahrscheinlichkeit w dafür, daß die Folge einem limes zwischen s_1 und s_2 zustrebt; w wächst gegen* 1, *wenn $n - m$ nach unendlich geht und dabei alle h_i $(m \leqq i \leqq n)$ zwischen diesen Schranken bleiben.*

Dieses Axiom spricht die Behauptung aus, daß ein limes existiert, und gibt zugleich an, wie sich der Wert des limes bestimmt. Es bestimmt den Wert des limes zwar nur mit einer gewissen Genauigkeit $\delta = s_2 - s_1$, aber mehr wird auch für praktische Anwendungen niemals verlangt. Das in dem Axiom auftretende Wort Wahrscheinlichkeit ist durch die in Zuordnungsdefinition α ausgesprochene Definition zu ersetzen. Ehe wir auf die eigenartige logische Struktur dieses Axioms eingehen, wollen wir zeigen, daß das Axiom für einige wichtige Schlüsse ausreicht, mit denen man es in Anwendungsproblemen zu tun hat.

Zunächst ergibt das Induktionsaxiom die sogenannte Wahrscheinlichkeitsbestimmung aposteriori, wenn man unter den Größen h_i speziell die Häufigkeiten $H_i(O, P)$ von Folgenpaaren versteht. Zugleich haben wir in dieser Form des Induktionsaxioms dasjenige Axiom vor uns, welches wir im vorangehenden Paragraphen aus Gründen der Sinndeutung der limes-

[24]) H. Reichenbach, Der Begriff der Wahrscheinlichkeit für die mathematische Darstellung der Wirklichkeit. Dissertation. Erlangen 1915, S. 63.

Forderung gefordert hatten. Das Axiom erlaubt dann nämlich, von einer beobachteten endlichen Wiederholungszahl rückwärts mit Wahrscheinlichkeit auf die Bestätigung der limes-Aussage zu schließen. — Für den Spezialfall $H_n = 1$ nimmt das Axiom die in der traditionellen Logik übliche Form an, wo aus dem wiederholten Eintreffen eines Ereignisses auf ständige Wiederholung geschlossen wird.

Eine weitere Anwendung des Axioms liegt vor, wenn man auf den Bernouillischen Charakter von Folgen nach (1, § 9) schließt. Hier sind die Größen h_i selbst schon limes-Größen, die durch vorangehende Anwendung des Induktionsaxioms im Sinne der Wahrscheinlichkeitsbestimmung aposteriori gewonnen worden sind.

Sodann erlaubt das Axiom, auf die Existenz stetiger Wahrscheinlichkeitsfunktionen zu schließen. Wir zeigen dies zunächst für die Wahrscheinlichkeitsfunktion *einer* Variablen. Man teilt dazu in dem betrachteten Abschnitt von a bis b (Fig. 2) die Abzisse in gleiche Intervalle $(x_1 - x_0)$, $(x_2 - x_1) \ldots (x_n - x_{n-1})$ ein, und ermittelt zunächst nach dem Verfahren der Wahrscheinlichkeitsbestimmung aposteriori, also unter erstmaliger Benutzung des Induktionsaxioms, die Wahrscheinlichkeit dafür, daß die betreffende Größe in das Intervall $(x_1 - x_0)$ bzw. $(x_2 - x_1)$ usw. fällt. Die so erhaltenen Wahrscheinlichkeiten trägt man in der Zeichnung derart ein, daß sie durch den Flächeninhalt des Ordinatenstreifens repräsentiert werden. Man erhält so eine Treppenkurve von endlich vielen Stufen, die eine maximale Differenz zweier aufeinanderfolgender Ordinaten bestimmen; wir nennen sie die „maximale Schwankung ε“. Das gleiche Verfahren wiederholt man nun mit einer feineren Intervallteilung, indem man also die Anzahl der Intervalle auf n_1 vergrößert; dabei vergrößert man auch die Zahl der Einzelbeobachtungen überhaupt, um für jedes Intervall etwa ebenso viele Treffer zu erhalten wie vorher. Man erhält auf diese Weise eine neue maximale Schwankung ε_1. Durch weitere Wiederholung des Verfahrens erhält man zwei gliedweise zugeordnete Folgen n_i und ε_i. Jetzt wendet man das Induktionsaxiom zum zweitenmal an, indem man die Größen ε_i als Größen h_i dieses Axioms auffaßt; die Erfahrung lehrt nun, daß die ε_i die Bedingung dieses Axioms derart erfüllen, daß das Axiom für die ε_i den limes 0 ergibt. Auf diese Weise ergibt sich, daß die Häufigkeit der Größen x_i im limes unendlichfacher Wiederholung durch eine stetige Wahrscheinlichkeitsfunktion bestimmt ist. Die Stetigkeit ist hier also nicht apriori vorausgesetzt, sondern mit Hilfe des Induktionsaxioms erschlossen. — Durch entsprechende Schlüsse

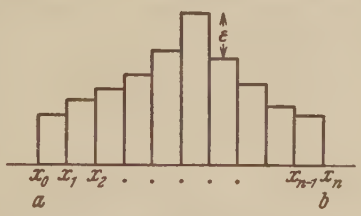


Fig. 2. Zur empirischen Ermittlung einer Wahrscheinlichkeitsfunktion.

ergibt sich das Bestehen stetiger Wahrscheinlichkeitsfunktionen für mehr als eine Variable.

Endlich wollen wir zeigen, daß das „Prinzip der einfachsten Kurve" aus unserem Induktionsaxiom folgt. Wir verstehen darunter das allgemein bei physikalischen Messungen benutzte Prinzip, eine Anzahl getrennter Messungspunkte durch eine möglichst einfache Kurve zu interpolieren (Fig. 3). Die übliche Darstellung, bei der die zu benutzten Werten x als zugehörig beobachteten Größen y aufgetragen werden, die wir auch in Fig. 3 angedeutet haben, läßt allerdings den Charakter der Kurve nicht deutlich hervortreten. Wir müssen uns vielmehr das Zustandekommen der Kurve folgendermaßen vorstellen. Die Koordinatenebene teilen wir durch ein quadratisches Netz, das Experiment liefert zu einem gewählten Abszissenintervall für jedes Netzquadrat des zugehörigen Ordinatenstreifens „Treffer", und schließlich ermitteln wir für jedes Netzquadrat die Wahrscheinlichkeit eines Treffers aposteriori, gerade wie in dem vorher behandelten Falle, wenn man ihn auf zwei Variable ausdehnt; durch Wiederholung des Verfahrens gehen wir dann wieder zum Grenzfall einer stetigen Wahrscheinlichkeitsfunktion der beiden Variablen x und y über. Diese Wahrscheinlichkeitsfunktion stellt die Wahrscheinlichkeit von x nach y dar (also das, was man in der bisherigen Terminologie die relative Wahrscheinlichkeit von y in bezug auf x nennt); im allgemeinen enthält das Vorderglied dieser Wahrscheinlichkeitsimplikation noch weitere Faktoren, die jedoch während der Variation von x konstant gehalten werden, so daß man mit einer Funktion von zwei Variablen auskommt. Die *Kammlinie dieses Wahrscheinlichkeitsgebirges* ist es, die man mit der den Messungspunkten angepaßten einfachsten Kurve meint. Daß man sich dieser Definition der einfachsten Kurve nicht bewußt ist, liegt in der üblichen Schematisierung physikalischer Feststellungen begründet; man pflegt dabei nämlich Wahrscheinlichkeitsfunktionen von hoher Präzision, die nur in einem schmalen Bereich von Null merklich verschieden sind, durch ihre Entartungsform zu ersetzen, welche die Wahrscheinlichkeit für eine einzige Argumentstelle in Gewißheit, für alle anderen Argumentstellen in Unmöglichkeit verwandelt. Diese für praktische Zwecke vielfach erlaubte, für erkenntnistheoretische Zwecke aber ganz unerlaubte Schematisierung führt auch zu der absolutistischen Auffassung der einfachsten Kurve als einer für sich allein bestehenden Figur; aber es ist klar, und wird übrigens auch durch das bei der wirklichen Herstellung benutzte Verfahren deutlich, bei welchem die Messungspunkte einen breiteren Streifen erfüllen, daß die

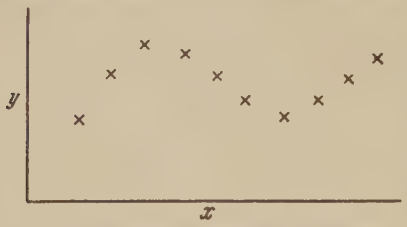


Fig. 3. Messungspunkte, die eine einfachste Kurve bestimmen.

einfachste Kurve nur als Kammlinie eines Wahrscheinlichkeitsgebirges, also als eingebettet in eine umfassendere Figur, definiert werden kann[25]). Bei der strengen Bestimmung der einfachsten Kurve, entsprechend ihrer Definition als Kammlinie, wird also das Induktionsaxiom in gleicher Weise benutzt wie bei der Bestimmung von Wahrscheinlichkeitsfunktionen mehrerer Veränderlicher überhaupt. Wenn man in praktischen Fällen die einfachste Kurve „durch die Messungspunkte hindurchzieht", so liegen dabei die beiden Annahmen zugrunde, daß erstens die eingezeichneten Messungspunkte bereits die Wahrscheinlichkeitsmaxima für ihren Ordinatenstreifen bedeuten, und daß zweitens die Maxima sich zu einer stetigen Kurve zusammenschließen, also die Wahrscheinlichkeitsfunktion stetig ist. Diese Annahmen müssen als Induktionsschlüsse aus früheren Erfahrungen betrachtet werden.

Fassen wir jetzt die Ergebnisse unserer Untersuchungen zur angewandten Wahrscheinlichkeitsrechnung für das Geltungsproblem zusammen, so haben wir festzustellen, daß ein Geltungsproblem nur für *ein* Axiom besteht, für das Induktionsaxiom; ein Axiom, welches innerhalb der mathematischen Wahrscheinlichkeitsrechnung überhaupt noch nicht auftritt. Die Axiome der mathematischen Wahrscheinlichkeitsrechnung sind in ihrer Anwendung auf die Wirklichkeit Tautologien, welche immer dann benutzt werden können, wenn Folgen von limes-Charakter vorliegen. Wenn man die Axiome der Wahrscheinlichkeitsrechnung immer als von eigentümlich zwingendem Charakter betrachtet hat, wenn man in ihnen „Grundsätze des vernunftgemäßen Schließens" gesehen hat, die denen des logischen Schließens analog sind, so klärt sich dieser Tatbestand durch unsere Untersuchung auf: man hat stets stillschweigend die Häufigkeitsdeutung der Wahrscheinlichkeit vorausgesetzt und damit die Axiome als Sätze rein logischer Natur benutzt, deren zwingender Charakter von derselben Art ist wie der aller andern logischen oder arithmetischen Sätze. Auch die speziellen Fälle der mathematischen Wahrscheinlichkeitsrechnung, normale Folgen und Bernoullische Folgen, enthalten kein Geltungsproblem; sie liegen immer dann vor, wenn gewisse Wahrscheinlichkeiten gleich werden bzw. einem limes zustreben. Der Entscheid darüber, ob dies vorliegt, wird aber wieder allein mit Hilfe des Induktionsaxioms getroffen. Es wäre ganz irrtümlich, in den Eigenschaften der normalen Folgen etwa allgemeine Naturgesetze zu sehen, deren Geltung besonderer philosophischer Begründung bedürfte. Es handelt sich hier um spezielle Gesetzlichkeiten, welche bei gewissen Naturvorgängen, keineswegs bei allen, gelten; darüber, ob dies der

[25]) Man kann noch hinzufügen, daß bei Verschärfung der Meßgenauigkeit das Wahrscheinlichkeitsgebirge sich immer mehr der Entartungsform annähert, die durch die Kammlinie gegeben ist; diese Annahme entspricht der üblichen Kausalhypothese, unterliegt aber deshalb genau wie diese empirischer Kritik und kann nicht apriori gefordert werden. Vgl. hierzu [5]).

Fall ist, entscheidet die Erfahrung. *Es gibt nur ein einziges erkenntnistheoretisches Prinzip, welches bei diesem Erfahrungsentscheid, ebenso wie in allen andern Anwendungsproblemen der Wahrscheinlichkeit vorausgesetzt wird; das ist das Axiom der Induktion.* In dieser Tatsache sehen wir eines der wichtigsten Ergebnisse der vorliegenden Untersuchung. Das Geltungsproblem der angewandten Wahrscheinlichkeitsrechnung reduziert sich deshalb auf das Problem der Geltung des Induktionsaxioms.

Was sich zur erkenntnistheoretischen Rechtfertigung dieses Axioms sagen läßt, habe ich an anderer Stelle ausgeführt[26]). Hier sei nur noch die Bemerkung ausgesprochen, daß ich nicht behaupten möchte, die oben gegebene Formulierung des Induktionsaxioms sei die allgemeinste, die überhaupt gegeben werden könne. Nach einem vorläufigen Überblick scheint sie auszureichen; dennoch ist es möglich, daß man, um alle in der Wissenschaft benutzten Schlußweisen zu rechtfertigen, die von uns gegebene Formulierung durch eine noch allgemeinere Formulierung ersetzen muß, welche die oben gegebene Formulierung als Spezialfall enthält.

§ 12.

Schlußbemerkung.

Wir müssen jetzt, nach der Aufstellung des Induktionsaxioms, noch einmal auf das Sinnproblem der Wahrscheinlichkeit zurückkommen. Wir hatten in § 10 ausgeführt, daß wir auf die Alternative wahr — falsch der strengen Logik verzichten und uns mit der Wahrscheinlichkeit von Aussagen begnügen wollen. Es entspricht diesem Gedanken, daß das Induktionsaxiom, welches den Grad p einer Wahrscheinlichkeit zu bestimmen lehrt, diesen Grad selbst nur mit einer gewissen Wahrscheinlichkeit w behauptet. Was aber überraschen mag, ist die Tatsache, daß diese Wahrscheinlichkeit zweiter Ordnung w dabei zunächst unbekannt ist.

Dies gilt jedoch nur zunächst; denn grundsätzlich ist w auf die gleiche Weise bestimmbar wie p, nämlich durch Auszählung einer Wiederholungsreihe, in der die einzelne Folge Element ist, unter nochmaliger Anwendung des Induktionsaxioms. Man erkennt übrigens leicht, daß die Wahrscheinlichkeit zweiter Ordnung damit in eine Beziehung zu der Dispersion der Folge tritt, und daß entsprechend die Wahrscheinlichkeiten höherer Ordnung durch die sogenannten Momente höheren Grades der Folge bestimmt sind. Trotz dieser grundsätzlichen zahlenmäßigen Bestimmbarkeit bleibt allerdings bestehen, daß jede tatsächliche Behauptung eines Wahrscheinlichkeitsgrades stets nur mit einer unbekannten Wahrscheinlichkeit ausgesprochen werden kann, die zwar zahlenmäßig bestimmbar ist, deren Bestimmung aber stets eine neue unbekannte Wahrscheinlichkeit an ihre Stelle

[26]) H. Reichenbach, Kausalität und Wahrscheinlichkeit, Erkenntnis I, 1930, S. 183.

setzt. Diese eigentümliche Doppelheit einer bestimmten und einer unbestimmten Wahrscheinlichkeit ist in jeder Wahrscheinlichkeitsaussage enthalten. Was man gelegentlich als Induktionsvertrauen bezeichnet, dürfte mit diesem Rest unbestimmter Wahrscheinlichkeit zu identifizieren sein; doch möchte ich diese Bezeichnung nicht übernehmen, da sie die Tatsache verdecken könnte, daß jede restliche unbestimmte Wahrscheinlichkeit grundsätzlich von derselben Art ist wie jede bestimmte Wahrscheinlichkeit, weil sie stets in eine bestimmte Wahrscheinlichkeit aufgelöst werden kann. In eben dieser Tatsache scheint mir die hinreichende Rechtfertigung enthalten zu sein, einen prinzipiellen Unterschied zwischen Wahrscheinlichkeit und Vermutung abzulehnen. Eine Vermutung ist nichts als die Behauptung einer Wahrscheinlichkeit ohne Angabe eines Wahrscheinlichkeitsgrades, der Unterschied also von (4, § 2) und (7, § 2). Daß eine Wahrscheinlichkeitsaussage stets diese beiden Formen der Wahrscheinlichkeitsimplikation enthält, bedeutet keine größere Schwierigkeit als der Übergang von der Wahrheitslogik zur Wahrscheinlichkeitslogik überhaupt.

Wir wollen zum Schluß die Eigenart der angewandten Wahrscheinlichkeitsrechnung noch einmal verdeutlichen, indem wir sie der inhaltlichen mathematischen Wahrscheinlichkeitsrechnung gegenüberstellen. In der inhaltlichen mathematischen Wahrscheinlichkeitsrechnung haben die Häufigkeitsaussagen den Charakter üblicher mathematischer Konvergenzaussagen, und die in dem Axiomensystem formulierten Zusammenhänge sind streng gültige mathematische Aussagen im üblichen Sinne. In der angewandten Wahrscheinlichkeitsrechnung sind die Axiome ebenfalls streng gültig, *wenn* die Wahrscheinlichkeitsfolgen den geforderten Konvergenzcharakter besitzen; aber die Aussage, *daß* die Folgen diesen Charakter besitzen, läßt sich weder durch strenge Logik begründen, noch besitzt sie den Sinn streng logischer Aussagen. Denn diese Aussage bedeutet die auf endlich viele Beobachtungen gestützte Behauptung einer Konvergenz für unendliche Verlängerung, einer Konvergenz zudem, deren Konvergenzstellen sich nicht angeben lassen; sie hat deshalb den Charakter einer *unbestimmten Prophezeiung* uud bedeutet damit eine Aussagenform, wie sie innerhalb der Mathematik niemals zugelassen werden kann. Die angewandte Wahrscheinlichkeitsrechnung ist deshalb keine mathematische Disziplin, und es ist grundsätzlich unmöglich, sie zu einer solchen zu machen. Da die Wahrscheinlichkeitsaussagen der Physik stets den Charakter von Aussagen der angewandten Wahrscheinlichkeitsrechnung tragen, ist deshalb die physikalische Wahrscheinlichkeitsaussage ebenfalls eine nicht-mathematisierbare Aussage; das Wahrscheinlichkeitsproblem der Physik kann nur im Rahmen der Wahrscheinlichkeitslogik seine Lösung finden.

(Eingegangen am 13. Mai 1931.)

Some new cases of Parseval's Theorem.

By

G. H. Hardy in Oxford and J. E. Littlewood in Cambridge.

1. Introduction.

1. In what follows we are concerned with a pair of analytic functions

$$f(z) = \sum_1^\infty a_n z^n, \quad g(z) = \sum_1^\infty b_n z^n, \tag{1.1}$$

regular for $r = |z| < 1$. It is convenient to suppose that the series have no constant terms. We call

$$h(z) = P(f, g) = \sum a_n b_n z^n \tag{1.2}$$

the "Faltung" or "Parseval function" of $f(z)$ and $g(z)$, and our problem is that of the convergence or summability of the series

$$P = \sum a_n b_n e^{ni\theta}, \tag{1.3}$$

the Faltung or Parseval series of the series

$$\sum a_n e^{ni\theta}, \quad \sum b_n e^{ni\theta}. \tag{1.4}$$

It is classical that P is absolutely convergent whenever f and g are L^2. More generally, suppose that

$$p > 1, \quad q = p' = \frac{p}{p-1}, \quad \frac{1}{p} + \frac{1}{q} = 1,$$

and that f and g are L^p and L^q respectively. Then it was proved by M. Riesz that P is convergent[1]), uniformly in θ, but in general not absolutely convergent. Riesz's theorem holds for general Fourier series, whereas we have stated it for power series only; but it follows from his theorem on conjugate functions that there is no real loss of generality in this limitation. Here we shall be concerned with cases in which p or q is less than 1, and the limitation to power series will prove to be essential.

[1]) That P is summable $(C, 1)$ had been proved earlier by Young (22). We have made extensions in a rather different direction; see Hardy and Littlewood (5

If $p<1$, q (as defined above) is negative, and Riesz's enunciation ceases to be significant. But there is a certain analogy between the Lebesgue class L^γ and the Lipschitz class $\mathrm{Lip}\left(-\frac{1}{\gamma}\right)$, a theorem true for functions of the first class when $\gamma>0$ being often also true, with appropriate changes, for functions of the second class when $\gamma<0$.[2]) Here this analogy suggests that, when f is L^p and $p<1$, the appropriate class for g is $\mathrm{Lip}\left(\frac{1}{p}-1\right)$. Guided by this analogy, we were led to the following theorems.

Theorem 1. *If*

(i) $0<k\leqq 1$, $pk\geqq 1$ (*so that* $\mu=\frac{p}{p+pk-1}\leqq 1$ *and* $p>1$ *if* $k<1$);

(ii) f *is* L^μ;

(iii) g *is* $\mathrm{Lip}(k,p)$ [3]);

then P *is uniformly convergent, and indeed uniformly summable* $\left(C,-\frac{1}{p}+\delta\right)$ *for every positive* δ.

Theorem 2. *If*

(i) $0<k\leqq 1$ $\left(\text{so that } \mu=\frac{1}{k+1}<1\right)$;

(ii) f *is* L^μ;

(iii) g *is* $\mathrm{Lip}\,k$;

then P *is uniformly summable* (C,δ) *for every positive* δ.

Theorem 3. *If, in addition to the conditions of Theorem* 1, $p\leqq 2$, *then* P *is absolutely convergent.*

It will be observed that Theorem 2 is the limiting case $p=\infty$ of Theorem 1. The question whether P is necessarily convergent in this case remains open. It is also to be observed that, in all the cases considered,

$$\mu\geqq\frac{1}{2}.$$

---

[2]) Witness for example Theorems 4 and 12 of our paper 7.

[3]) It may be convenient that we should repeat the definitions of the classes $\mathrm{Lip}\,k$ and $\mathrm{Lip}(k,p)$. The function $g(z)$ belongs to $\mathrm{Lip}\,k$ if

$$|g(z')-g(z)|\leqq C|z'-z|^k,$$

where C is independent of z and z', and to the class $\mathrm{Lip}(k,p)$ if

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}|g(re^{i\theta+ih})-g(re^{i\theta-ih})|^p\,d\theta\leqq C|h|^{pk},$$

where C is independent of r.

There are doubtless similar theorems in which $\mu < \frac{1}{2}$, but they will plainly demand definitions of Lipschitz classes of order greater than 1 (in which there is of course no difficulty of principle).

Our theorems are largely applications of theorems which we have proved in previous papers, and in particular in **20** (to which, and to the earlier papers of this series, we shall have to refer continually).

2. Convergence and summability: proof of Theorems 1 and 2.

2.1. In this section we shall make free use of theorems from **20** and one preliminary remark is required. There, we defined the fractional integral $f_\alpha(z)$ by the formula

$$f_\alpha(z) = \sum \frac{\Gamma(n+1)}{\Gamma(n+1+\alpha)} a_n z^{n+\alpha}. \tag{2.11}$$

Here, we shall write instead

$$f_\alpha(z) = \sum_1^\infty n^{-\alpha} a_n z^n. \tag{2.12}$$

The reader will find no difficulty in satisfying himself that any property quoted for the one form of f_α is also possessed by the other.

Suppose, for example, that one of the functions belongs to L^p. The factor z^α in (2.11) is unimportant; and apart from this factor the two forms differ only by a factor of the type

$$1 + \frac{c_1}{n+1} + \frac{c_2}{(n+1)(n+2)} + \dots \text{ [4]}$$

in the coefficient of z^n. If then we call the two functions φ and ψ, ψ is the sum of φ, a finite linear combination of ordinary integrals of φ, and a continuous function. By Theorem 33 of **20**, the integrals of φ belong to L^p (and indeed to higher classes). Hence, if φ belongs to L^p, ψ also belongs to L^p.

The advantage of the definition (2.12), for our present purpose, is that

$$h = P(f, g) = P(f_\alpha, g^\alpha) \tag{2.13}$$

for any real α.

2.2. Passing to the proof of Theorem 1, we dispose first of the case $p = 1$, $k = 1$, $\mu = 1$. In this case, if $f(e^{i\theta})$ and $g(e^{i\theta})$ are the boundary functions of $f(z)$ and $g(z)$, the first is integrable and the second (by Theorem 24 of 7) is of bounded variation; and

$$h(e^{i\theta}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{i\varphi})\, g(e^{i\theta - i\varphi})\, d\varphi$$

[4]) A finite series followed by an error term of lower order than any preceding term. If $f(z)$ belongs to any L^p, then $a_n = O(n^K)$ for some K.

s also of bounded variation. It follows, by Theorem 16 of 4, that $\sum |a_n b_n|$ s convergent. Finally, since $a_n b_n = O\left(\frac{1}{n}\right)$, the series (1.3) is summable $(C, -1+\delta)$ for every positive δ.

We may then suppose that $p > 1$. We prove that

(i) P is the Fourier (power) series of a continuous function,

(ii) P is the Fourier (power) series of a function of the class

$$\mathrm{Lip}\left(\frac{1}{p}, p\right)$$

or A_p. It will follow from (i) that P is uniformly summable $(C, 1)$, and from (ii), by Theorem 1 of 8 [5]), that it is uniformly summable $\left(C, -\frac{1}{p}+\delta\right)$ and in particular uniformly convergent.

2.3. Proof of (i). We shall prove rather more, viz. that

$$\int_0^1 |h'(re^{i\theta})|\, dr$$

s uniformly convergent. *A fortiori*, $h(e^{i\theta})$ is continuous and P is its Fourier series.

If $r = \varrho^2 < 1$, we have

$$r h'(re^{i\theta}) = \sum n a_n b_n \varrho^{2n} e^{ni\theta} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\varrho e^{i\varphi})\, g'(\varrho e^{i\theta - i\varphi})\, d\varphi,$$

$$|r h'(re^{i\theta})| \leqq M_{p'}(f)\, M_p(g'). \text{ [6])}$$

By Theorem 48 of 20 (Theorem 3 of 8),

$$M_p(g') = O\left((1-\varrho)^{k-1}\right),$$

and our result will plainly follow if

$$\int_0^1 (1-\varrho)^{k-1} M_{p'}(f)\, d\varrho$$

s convergent. This follows from Theorem 31 of 20, when we replace α by μ, q by p', and l by 1, so that $l\alpha$ is replaced by

$$\frac{1}{\mu} - \frac{1}{p'} = 1 + k - \frac{1}{p} - \frac{1}{p'} = k.$$

---

[5]) There is no reference to uniformity in 8, where we are concerned with a particular point; but the whole argument is "uniform in θ", the difference between any two Cesàro means, of order greater than $-\frac{1}{p}$, being shown to tend uniformly to zero.

[6]) As usual $p' = \frac{p}{p-1}$.

2.4. Proof of (ii). By Theorem 33 of **20**, $f_{k-\frac{1}{p}}(z)$ belongs to L (since

$$\frac{\mu}{1-\left(k-\frac{1}{p}\right)\mu}=1);$$

and, by Theorem 49, $g^{k-\frac{1}{p}}(z)$ belongs to $\mathrm{Lip}\left(k-k+\frac{1}{p}, p\right)$ or Λ_p. Also h is the Faltung of $f_{k-\frac{1}{p}}$ and $g^{k-\frac{1}{p}}$. It is therefore sufficient to prove that the Faltung of a function F of L and a function G of Λ_p belongs to Λ_p.

But in these circumstances, writing $\Delta F(\theta)$ for $F(\theta+h)-F(\theta-h)$, and supposing $h>0$, we have

$$|\Delta H(\theta)| \leqq \frac{1}{2\pi}\int_{-\pi}^{\pi}|F(\varphi)|\,|\Delta G(\theta-\varphi)|\,d\varphi,$$

$$\left(\int|\Delta H(\theta)|^p\,d\theta\right)^{\frac{1}{p}} \leqq \frac{1}{2\pi}\int|F(\varphi)|\,d\varphi\left(\int|\Delta G(\theta-\varphi)|^p\,d\theta\right)^{\frac{1}{p}} = O\left(h^{\frac{1}{p}}\right),$$

the result required. This completes the proof of Theorem 1.

2.5. Proof of Theorem 2. The argument of § 2.3 is valid as it stands in the limiting case $p=\infty$, if we replace $M_{p'}$ and M_p by M_1 and M_∞ or M. This is all that is required to prove the theorem.

The argument of § 2.4 fails, and we cannot assert convergence; but we have no 'Gegenbeispiel'.

3. Absolute convergence: proof of Theorem 3.

3. We shall give two different proofs of Theorem 3. The first, in which we make use of theorems from **20** and of what we have proved already, is much the shorter; but the second depends on subsidiary propositions which are interesting in themselves, and also proves more.

(i) We observe first that it is sufficient to prove the theorem in the case $p=2$. For suppose that g belongs to $\mathrm{Lip}(k,p)$ and that $p<q$. Then, by Theorem 5 of 8, g also belongs to $\mathrm{Lip}(l,q)$, where

$$l=k-\frac{1}{p}+\frac{1}{q};$$

and

$$\mu(l,q)=\frac{q}{q+ql-1}=\frac{p}{p+pk-1}=\mu(k,p).$$

It follows that, if Theorem 3 is true for q, l, μ, it is also true for p, k, μ. The same argument would show, of course, that Theorem 1 is true for a given p if it is true for some larger p.

(ii) We may therefore suppose that $p = 2$, when we have to prove that $\sum |a_n b_n|$ is convergent for an f and g belonging respectively to

$$L^{\frac{2}{1+2k}}, \quad \mathrm{Lip}\,(k, 2).$$

The hypothesis on g, viz.

$$\int |g(e^{i\theta+ih}) - g(e^{i\theta-ih})|^2 d\theta = O(h^{2k}),$$

or

$$\sum |b_n|^2 \sin^2 nh = O(h^{2k}),$$

is then equivalent to

$$\sum_1^n |\nu b_\nu|^2 = O(n^{2-2k}),$$

and is a hypothesis on the moduli $|b_n|$ only; the arguments $\arg b_n$ are at our disposal. It follows that $\sum a_n b_n$, which we know already to be convergent, is absolutely convergent.

Our second proof demands some preliminary theorems concerning series of positive terms. These have some interest in themselves, and we have developed them a little further than is necessary for our application.

4. Some theorems on series of positive terms.

4.1. Theorem 4. *If x_n, y_n, and λ_n are positive;*

$$p > 1, \quad \frac{1}{p} + \frac{1}{q} = 1, \quad r > 1;$$

and

$$\sum x_n^p < \infty, \quad \sum y_n^q < \infty, \quad \sum_1^n \lambda_\nu^r = O(n);$$

then

$$\sum\sum \frac{x_m y_n}{m+n} \lambda_{m+n} < \infty.$$

The theorem becomes false if we replace r by 1.

The summations are over $m, n = 1, 2, 3, \ldots$. The case $\lambda_n = 1$, $p = q = 2$ is Hilbert's double series theorem.

We write

$$\Lambda_n = \lambda_1 + \lambda_2 + \ldots + \lambda_n,$$

so that

$$\Lambda_n \leqq n^{\frac{1}{r'}} \left(\sum_1^n \lambda_\nu^r\right)^{\frac{1}{r}} = O(n).$$

We have[7])

$$S = \sum\sum \frac{x_m y_n}{m+n} \lambda_{m+n} = \sum\sum x_m \left(\frac{\lambda_{m+n}}{m+n}\right)^{\frac{1}{p}} \left(\frac{m}{n}\right)^s \cdot y_n \left(\frac{\lambda_{m+n}}{m+n}\right)^{\frac{1}{q}} \left(\frac{n}{m}\right)^s,$$

[7]) Compare Hardy, Littlewood, and Pólya (21, p. 280).

where s is a positive number to be chosen later. By Hölder's inequality

$$S \leqq P^{\frac{1}{p}} Q^{\frac{1}{q}},$$

where

$$P = \sum\sum x_m^p \left(\frac{m}{n}\right)^{sp} \frac{\lambda_{m+n}}{m+n}, \qquad Q = \sum\sum y_n^q \left(\frac{n}{m}\right)^{sq} \frac{\lambda_{m+n}}{m+n};$$

and it is enough to prove the convergence of P and Q.

Now

$$P = \sum_m x_m^p m^{sp} \sum_n n^{-sp} \frac{\lambda_{m+n}}{m+n} = \sum x_m^p m^{sp} T_m, \tag{4.11}$$

say. Summing partially, and observing that $\Lambda_n = O(n)$, we obtain

$$\begin{aligned} T_m &= \sum_n \frac{n^{-sp}}{m+n} \{(\Lambda_{m+n} - \Lambda_m) - (\Lambda_{m+n-1} - \Lambda_m)\} \\ &= \sum_n (\Lambda_{m+n} - \Lambda_m) \Delta \frac{n^{-sp}}{m+n}, \end{aligned}$$

where Δu_n denotes $u_n - u_{n+1}$. Now

$$\Delta \frac{n^{-sp}}{m+n} < K \frac{n^{-sp-1}}{m+n},$$

where K is independent of m and n. Hence

$$T_m \leqq K \sum_n (\Lambda_{m+n} - \Lambda_n) \frac{n^{-sp-1}}{m+n}.$$

But

$$\Lambda_{m+n} - \Lambda_m = \sum_{m+1}^{m+n} \lambda_\nu \leqq n^{\frac{1}{r'}} \left(\sum_{m+1}^{m+n} \lambda_\nu^r\right)^{\frac{1}{r}} < K n^{\frac{1}{r'}} (m+n)^{\frac{1}{r}};$$

and so

$$T_m \leqq K \sum_n n^{\frac{1}{r'} - sp - 1} (m+n)^{\frac{1}{r} - 1} = O(m^{-sp}),$$

provided only that $0 < sp < \frac{1}{r'}$. If s satisfies this condition, it follows from (4.11) that P is convergent.

Similarly Q is convergent if $0 < sq < \frac{1}{r'}$. Since we can choose s to satisfy both these conditions, the theorem is proved.

To prove that the result is false when $r = 1$, take $\lambda_n = n$ when $n = 2$ and $\lambda_n = 0$ otherwise. The theorem would then assert that

$$\sum_\nu \sum_{m+n=2^\nu} x_m y_n$$

is convergent whenever $\sum x_m^p$ and $\sum y_n^q$ are convergent, and this is false

4.2. Theorem 5. *If*

$$p>1, \quad X_n = x_1 + x_2 + \ldots + x_n,$$

and

$$\sum x_n^p < \infty, \quad \Lambda_n = O(n),$$

then

$$\sum \left(\frac{X_n}{n}\right)^p \lambda_n < \infty.$$

We do not require this theorem, but we insert it because it corresponds to a theorem of Hardy[8]) as Theorem 4 corresponds to Hilbert's theorem. It will be observed that a difference between the two theorems appears when (in the notation of Theorem 4) $r=1$.

Suppose that $\Lambda_n \leqq Cn$. Then

$$\sum_1^N \left(\frac{X_n}{n}\right)^p \lambda_n = \sum_1^{N-1} \Lambda_n \Delta\left(\frac{X_n}{n}\right)^p + \Lambda_N \left(\frac{X_N}{N}\right)^p,$$

and the last term tends to zero because $\Lambda_N = O(N)$ and $X_N = o\left(N^{\frac{1}{p'}}\right)$.[8a]) Hence

$$S = \sum \left(\frac{X_n}{n}\right)^p \lambda_n = \sum \Lambda_n \Delta\left(\frac{X_n}{n}\right)^p \leqq CK \sum \left(\frac{X_n}{n}\right)^p + CK \sum' \left(\frac{X_n}{n}\right)^{p-1} x_n$$

$$\leqq CK \sum \left(\frac{X_n}{n}\right)^p + CK \left(\sum' \left(\frac{X_n}{n}\right)^p\right)^{\frac{1}{p'}} \left(\sum x_n^p\right)^{\frac{1}{p}} \leqq CK \sum x_n^p,$$

the K being positive numbers which depend only on p.

4.3. Theorem 6. *If* $p>1$,

$$\chi_n = \sum_m e^{-\frac{m}{n}} x_m,$$

and

$$\sum x_n^p < \infty, \quad \Lambda_n = O(n),$$

then

$$\sum \left(\frac{\chi_n}{n}\right)^p \lambda_n < \infty.$$

This theorem is related to one of the theorems[9]) of our paper **19** (to which it reduces when $\lambda_n = 1$), as Theorems 4 and 5 are to the other known theorems to which we have referred.

We suppose again that $\Lambda_n \leqq Cn$. We can choose M so that

$$\sum_{M+1}^{\infty} x_m^p < \varepsilon^p,$$

---

[8]) $\sum \left(\frac{X_n}{n}\right)^p < \infty$.

[8a]) Compare the proof in § 4.3 of the corresponding property of χ_n.

[9]) Theorem 2 (2.54).

and then

$$\chi_n \leqq \sum_1^M e^{-\frac{m}{n}} x_m + e\left(\sum_{M+1}^{\infty} e^{-\frac{p'm}{n}}\right)^{\frac{1}{p'}};$$

from which it follows that $\chi_n = o\left(n^{\frac{1}{p'}}\right)$. But

$$\sum_1^N \left(\frac{\chi_n}{n}\right)^p \lambda_n = \sum_1^{N-1} \Lambda_n \Delta\left(\frac{\chi_n}{n}\right)^p + \Lambda_N \left(\frac{\chi_N}{N}\right)^p,$$

and the last term tends to zero. Hence

$$S = \sum \left(\frac{\chi_n}{n}\right)^p \lambda_n = \sum \Lambda_n \Delta\left(\frac{\chi_n}{n}\right)^p \leqq CKS_1 + CKS_2,$$

where

$$S_1 = \sum \left(\frac{\chi_n}{n}\right)^p, \quad S_2 = \sum \left(\frac{\chi_n}{n}\right)^{p-1} |\Delta\chi_n|^{10}),$$

and the K depend only on p. Since $S_1 < \infty$ (by the theorem of 19 just referred to), and

$$S_2 \leqq \left(\sum \left(\frac{\chi_n}{n}\right)^p\right)^{\frac{1}{p'}} \left(\sum |\Delta\chi_n|^p\right)^{\frac{1}{p}},$$

it is sufficient to prove that

$$T = \sum |\Delta\chi_n|^p < \infty.$$

But

$$|\Delta\chi_n| = \sum_m \left(e^{-\frac{m}{n+1}} - e^{-\frac{m}{n}}\right) x_m \leqq \frac{4}{(2n)^2} \sum m e^{-\frac{m}{2n}} x_m \leqq \frac{4}{(2n)^2} \sum K n e^{-\frac{m}{4n}} x_m = K\frac{\chi_{4n}}{4n}$$

and so

$$T = \sum |\Delta\chi_n|^p \leqq K \sum \left(\frac{\chi_{4n}}{4n}\right)^p \leqq K \sum \left(\frac{\chi_n}{n}\right)^p \leqq K \sum x_m^p.$$

This completes the proof of the theorem.

5. Direct proof of Theorem 3.

5.1. We return to Theorem 3, proving, as we said in § 3, rather more.

Theorem 7. *If* $\mu < 1$ *and*

(i) $f(z)$ *is* L^μ,

(ii) $|c_1| + |c_2| + \ldots + |c_n| = O(n)$,

then

$$\sum n^{\mu-2} |a_n|^\mu |c_n|$$

is convergent.

---

[10]) $\Delta\chi_n$ is negative.

When $c_n = 1$, this reduces to Theorem 16 of 4 [11]).

We may plainly suppose that c_n is positive. It is also sufficient on grounds to which we have appealed many times in this series of apers, and need not repeat again) to prove the theorem for an $f(z)$ hich has no zeros in the unit circle.

We write

$$k = \frac{2}{\mu} > 2, \quad r = 2 - \mu = \frac{2}{k'} > 1$$

so that $(k-1)\mu = r$), and

$$f = \varphi^k, \quad \varphi = \sum u_n z^n$$

so that φ belongs to L^2). Let

$$\varrho = e^{-\frac{1}{k'n}}.$$

hen

$$|a_n| \leqq \frac{\varrho^{-n}}{2\pi}\int_{-\pi}^{\pi} |f(\varrho e^{i\theta})|\,d\theta \leqq K\int_{-\pi}^{\pi} |\varphi|^k\,d\theta \leqq K\Big(\sum (|u_m|\varrho^m)^{k'}\Big)^{k-1},$$

y Hausdorff's theorem, K depending only on k. If now we write $_n = |u_m|^{k'}$, $\lambda_m = |c_m|$, and define χ_n as in Theorem 6, we have

$$|a_n| \leqq K\chi_n^{k-1}$$

nd

$$\sum n^{\mu-2}|a_n|^\mu |c_n| \leqq K\sum \chi_n^{(k-1)\mu} n^{\mu-2}\lambda_n = K\sum \Big(\frac{\chi_n}{n}\Big)^r \lambda_n.$$

ut

$$\sum x_n^r = \sum |u_n|^2$$

convergent, since φ is L^2; and our conclusion therefore follows from heorem 6.

Theorem 8. *If the conditions of Theorem 7 are satisfied, then*

$$\sum n^{-\frac{1}{\mu}} |a_n|\,|c_n|$$

convergent.

This is included in Theorem 7, since

$$a_n = o\Big(n^{\frac{1}{\mu}-1}\Big),$$

Theorem 28 of 20.

5. 2. When $\mu < 1$ (*i. e.* when $pk > 1$) Theorem 3 may be deduced om Theorem 8. For, by Hausdorff's theorem and Theorem 48 of 20,

$$\Big(\sum (|nb_n|r^n)^{p'}\Big)^{\frac{1}{p'}} \leqq M_p(g') = O((1-r)^{k-1}).$$

[11]) Stated as a convergence theorem.

Hence, taking $r = e^{-\frac{1}{N}}$, we see that

$$\sum_1^N |n b_n|^{p'} = O(N^{p'(1-k)}),$$

$$\sum_1^N |n^{\frac{1}{\mu}} b_n|^{p'} = O\left(N^{p'\left(\frac{1}{\mu}-k\right)}\right) = O(N),$$

and *a fortiori*

$$\sum_1^N |n^{\frac{1}{\mu}} b_n| = O(N).$$

We may therefore take $c_n = n^{\frac{1}{\mu}} b_n$ in Theorem 8, and this proves Theorem 3.

5.3. Theorems 7 and 8 are false when $\mu = 1$. In fact, in the notation of § 5.1, $k = 2$, $f = \varphi^2$, and

$$\sum_n n^{-1} |a_n| \, |c_n| = \sum_n n^{-1} |c_n| \, \Big| \sum_{\mu+\nu=n} u_\mu u_\nu \Big|;$$

and it is plain (since the condition on f or φ is a condition on the moduli $|u_n|$ only) that each of the theorems is equivalent to the forbidden case $r = 1$ in Theorem 4.

Theorem 3 is however still true, since the c_n of § 5.2 satisfies

$$\sum_1^N |c_n|^r = O(N)$$

for $r = p' > 1$, so that the theorem follows from Theorem 4.

5.4. We shall not attempt to show systematically that the theorems which we have proved are the 'best possible' of their kind, but we shall discuss shortly some examples whose general tendency is in this direction.

(i) We stated in § 1 that the limitation of our theorems to power series is essential. To illustrate this, consider the analogue for general Fourier series of the case $pk = 1$, $\mu = 1$ of Theorem 1. This would imply that the Faltung of any two real functions $F(\theta)$ and $G(\theta)$, of which F is L and G is Λ_p, is continuous; and this would imply that any G of Λ_p is bounded. That this is untrue is shown by the example

$$G(\theta) = \log \frac{1}{|\theta|},$$

which is Λ_p for every $p > 1$.

The same point may be illustrated rather differently. Take

$$F(r, \theta) = 1 + 2\sum r^n \cos n\theta, \qquad G(r, \theta) = \sum \frac{r^n}{n} \cos n\theta.$$

Then F is L, since $M_1(F) = 1$, and G is Λ_p, but the Parseval series diverges for $\theta = 0$.

If we desire similar examples with $\mu < 1$, it is natural to use the function

$$F(r,\theta) = \left(\frac{d}{d\theta}\right)^{\beta}\left(1 + 2\sum r^n \cos n\theta\right) = 2\sum n^{\beta} r^n \cos\left(n\theta + \frac{1}{2}n\beta\right),$$

where β is a positive integer. It is easily verified that F is $L^{\frac{1}{1+\beta}}$. In this case, however, $\mu \leqq \frac{1}{2}$, so that either $k > 1$ (when the example is not relevant here) or $k = 1$, $p = \infty$ (in which case it is used in (2) below).

(2) In Theorem 2 we left the question of *convergence* unsettled. The function just mentioned, with $\beta = 1$, enables us to answer the corresponding question for *harmonic* functions negatively. Take

$$F(r,\theta) = \sum n r^n \sin n\theta,$$

and for $G(\theta)$ any odd function of Lip 1 (*i. e.* any odd function which is the integral of a bounded function). Then $k = 1$, $p = \infty$, $\mu = \frac{1}{2}$, and the conditions corresponding to those of Theorem 2 are satisfied. It is not true that $\sum n b_n$ is necessarily convergent (since the Fourier series of a bounded function is not necessarily convergent).

(3) The example

$$f(z) = \sum \frac{n^{\beta} z^n}{(\log n)^{\gamma}}, \qquad g(z) = \sum \frac{e^{ai n \log n}}{n^{k+\frac{1}{2}}} z^n,$$

where

$$a > 0, \quad 0 < k < 1, \quad p > 2, \quad \beta = k - \frac{1}{p}, \quad \gamma > \frac{1}{\mu},$$

shows that Theorem 3 is false when $p > 2$. Here f is L^{μ} and g is Lip k, *a fortiori* Lip (k, p); and $\sum |a_n b_n|$ is divergent.

(4) In Theorems 1—3 we have always $pk \geqq 1$ and so $\mu \leqq 1$. If $pk < 1$ then $\mu > 1$; and we give an example to illustrate the failure of the theorems in this case. Suppose that $0 < p < 2$ and $k = \frac{1}{p} - \frac{1}{2}$, so that $\mu = 2$, and take $g(z) = (1-z)^{-\frac{1}{2}}$; it is easily verified that g is Lip (k, p). If the conclusion of Theorem 1 were true, it would follow that the convergence of $\sum |a_n|^2$ implies that of $\sum n^{-\frac{1}{2}} a_n$; and this is false.

6. On a theorem of W. H. Young.

6.1. The theorems of this and the next section are of a different type, but also depend on theorems from **20**.

It was proved by Young[12]) that if

(6.11) $$p \geqq 1, \quad q \geqq 1, \quad \frac{1}{p} + \frac{1}{q} > 1,$$

$F(\theta)$ and $G(\theta)$ are any functions of L^p and L^q, and $H(\theta)$ is their Parseval function or Fourier Faltung, then $H(\theta)$ belongs to L^{μ}, where

(6.12) $$\mu = \frac{pq}{p+q-pq}.$$

[12]) Young (**23**).

In particular, when $F(\theta)$ and $G(\theta)$ are the boundary functions $f(e^{i\theta})$ and $g(e^{i\theta})$, we have

Theorem 9. *If p and q satisfy* (6.11), *$f(z)$ is L^p, and $g(z)$ is L^q, then $h(z)$ is L^μ and*

$$M_\mu(h) \leqq M_p(f)\, M_q(g). \tag{6.13}$$

The question arises whether this theorem for power series remains true when p or q is less than 1, and, if not, what replaces it. The answer to the first question is easily seen to be negative; if

$$f(z) = \sum_2^\infty \frac{n^\alpha}{(\log n)^\gamma} z^n,$$

where α and γ are positive and $\frac{\gamma}{1+\alpha} > 1$, and

$$g(z) = \sum 2^{-\beta n} z^{2^n},$$

where $0 < \beta < \alpha$, then f belongs to $L^{\frac{1}{1+\alpha}}$ and g is continuous, but

$$h(z) = \frac{1}{(\log 2)^\gamma} \sum \frac{2^{n(\alpha-\beta)}}{n^\gamma} z^{2^n}$$

does not belong to any Lebesgue class.

It is known that if h belongs to L^μ, where $\mu > 0$, then

$$\int_0^1 |h|^\mu \, dr \tag{6.14}$$

is (uniformly) convergent[13]), and that, if $0 < \mu \leqq 2$, and

$$\mathfrak{H}(r) = \sum |a_n|\,|b_n|\, r^n,$$

then

$$\int_0^1 \mathfrak{H}^\mu(r)\, dr \tag{6.15}$$

is convergent[14]), the converse propositions being false. This suggests that the appropriate conclusion, when we allow p or q to be less than 1 in Young's theorem, may be the convergence of (6.14) or (6.15).

Theorem 10. *If*

$$p > 0, \quad q > 0, \quad \frac{1}{p} + \frac{1}{q} > 1,$$

f is L^p and g is L^q, then (6.14) *is (uniformly) convergent.*

Theorem 11. *If, in addition to the conditions of Theorem* 10, *$p \leqq 2$, $q \leqq 2$, then* (6.15) *is convergent.*

To prove Theorem 10 we observe that, since $\frac{1}{p} + \frac{1}{q} > 1$, we can choose k so that

$$k > p, \quad k' > q, \quad \frac{1}{k} + \frac{1}{k'} = 1.$$

[13]) Fejér and F. Riesz (18).

[14]) Hardy and Littlewood (4, Theorem 15).

Then, if $r = \varrho^2 < 1$, we have

$$(6.16)\quad |h(re^{i\theta})| = \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} f(\varrho e^{i\varphi})\, g(\varrho e^{i\theta - i\varphi})\, d\varphi\right| \leqq M_k(\varrho, f)\, M_{k'}(\varrho, g),$$

$$\int_0^1 |h(re^{i\theta})|^\mu d\varrho \leqq \int_0^1 M_k^\mu(f)\, M_{k'}^\mu(g)\, d\varrho$$

$$\leqq \left(\int_0^1 M_k^{\frac{pk}{k-p}}(f)\, d\varrho\right)^{\frac{(k-p)\mu}{pk}} \left(\int_0^1 M_{k'}^{\frac{qk'}{k'-q}}(g)\, d\varrho\right)^{\frac{(k'-q)\mu}{qk'}} < \infty,$$

by Theorem 31 of **20**, provided only that $k > p$, $k' > q$, and

$$\frac{(k-p)\mu}{pk} + \frac{(k'-q)\mu}{qk'} = 1,$$

all of which conditions are satisfied.

If $p < 2$, $q < 2$ we can suppose that $k = 2$, and then we can replace (6.16) by

$$\mathfrak{H}(r) = \Sigma |a_n|\,|b_n|\, \varrho^{2n} \leqq M_2(\varrho, f)\, M_2(\varrho, g).$$

We then obtain Theorem 11 by the same argument as before. The case in which p or q is 2 requires a trivial modification of the argument which may be left to the reader.

If $p \geqq 1$, $q \geqq 1$, these theorems are naturally corollaries of Theorem 9, in virtue of the theorems of Fejér-Riesz and Hardy-Littlewood just referred to.

References.

(For numbers less than 18 see the list of papers at the end of **20**.)

18. L. Fejér und F. Riesz, Über einige funktiontheoretische Ungleichungen, Math. Zeitschr. **11** (1921), S. 305—314.
19. G. H. Hardy and J. E. Littlewood, Elementary theorems concerning power series with positive coefficients and moment constants of positive functions, Journ. für Math. **157** (1927), S. 141—158.
20. — — Some properties of fractional integrals (II), Math. Zeitschrift **34** (1931), S. 403—439.
21. — — G. Pólya, The maximum of a certain bilinear form, Proc. London Math. Soc. (2) **25** (1926), p. 265—282.
22. W. H. Young, On a class of parametric integrals and their application in the theory of Fourier series, Proc. Roy. Soc. (A), **85** (1911), p. 401—414.
23. — On the multiplication of successions of Fourier constants, ibidem **87** (1912), p. 331—339.

(Eingegangen am 27. April 1931.)

An additional note on Parseval's Theorem.

By

G. H. Hardy and J. E. Littlewood in Cambridge.

1. In our paper "Some new cases of Parseval's Theorem"[1]) we left open a question to which we have since found the answer.

Suppose that $0 < \alpha \leqq 1$, that

$$f(z) = \sum_1^\infty a_n z^n, \qquad g(z) = \sum_1^\infty b_n z^n,$$

and that $f(z)$ and $g(z)$ belong to the classes

$$L^{\frac{1}{1+\alpha}}, \qquad \mathrm{Lip}\,\alpha$$

respectively. Is it true that $\sum a_n b_n$ is necessarily convergent? The answer is (as was to be expected) negative.

2. We shall prove (what is, except when $\alpha = 1$, rather more) that $\sum a_n b_n$ is not necessarily convergent when f belongs to $L^{\frac{1}{1+\alpha}}$ and g is the α-th integral of a bounded function. Since the Parseval series of f and g is identical with that of f_α and g^α, it is the same thing to prove that the proposition

(A) "$\sum a_n b_n$ is convergent whenever f^α belongs to $L^{\frac{1}{1+\alpha}}$ and g is bounded"

is false.

The "inequality" form of (A) would be

$$\left|\sum_1^N a_n b_n\right| \leqq K(\alpha)\,\mathrm{Max}\,|g|\left(\int_{-\pi}^{\pi} |f^\alpha|^{\frac{1}{1+\alpha}} d\theta\right)^{1+\alpha}, \tag{1}$$

and we begin by disproving this. The disproof of (A) itself is then a matter of familiar routine.

[1]) Math. Zeitschr. **34** (1931), S. 620–633. We refer to this paper as P.

Suppose first that $\alpha = 1$. It is known [2]) that there are functions g for which $|g| \leqq 1$ and

$$|B_N| = |b_1 + b_2 + \ldots + b_N| > K \log N \tag{2}$$

for a constant positive K and an infinity of N. We choose such a g and such an N, and define $f(z) = f_N(z)$ by

$$f^1(z) = a_1 z + 2a_2 z^2 + 3a_3 z^3 + \ldots = \frac{z}{N+1}\left(\frac{1-z^{N+1}}{1-z}\right)^3. \tag{3}$$ [3])

Then, writing ν for $N+1$, we have

$$\int |f^1|^{\frac{1}{2}} d\theta \leqq \frac{1}{\sqrt{\nu}} \int_{-\pi}^{\pi} \left| \frac{\sin \frac{1}{2}\nu\theta}{\sin \frac{1}{2}\theta} \right|^{\frac{3}{2}} d\theta \leqq \frac{K}{\sqrt{\nu}} \int_0^{\frac{1}{\nu}} \nu^{\frac{3}{2}} d\theta + \frac{K}{\sqrt{\nu}} \int_{\frac{1}{\nu}}^{\infty} \theta^{-\frac{3}{2}} d\theta < K. \tag{4}$$

On the other hand, the first N terms of $f(z)$ are

$$\frac{1}{2(N+1)} \sum_1^N (n+1) z^n,$$

and so

$$\sum_1^N a_n b_n = \frac{2b_1 + 3b_2 + \ldots + (N+1) b_N}{2(N+1)} \tag{5}$$

$$= \frac{1}{2} B_N - \frac{B_1 + B_2 + \ldots + B_{N-1}}{2(N+1)} > K \log N,$$

since $B_N > K \log N$, by (2), and the second term, which is substantially a Cesàro mean formed from the bounded function g, is bounded. It follows from (4) and (5) that (1) is false when $\alpha = 1$ and f and g are chosen suitably.

The same example shows the falsity of (1) for general α, since

$$\left(\int |f^{\alpha}|^{\frac{1}{1+\alpha}} d\theta \right)^{1+\alpha} \leqq K(\alpha) \left(\int |f^1|^{\frac{1}{2}} d\theta \right)^2.$$ [4])

3. The falsity of (A) itself may now be proved as follows. We suppose for simplicity of writing that $\alpha = 1$.

---

[2]) H. Bohr, „Über die Koeffizientensumme einer beschränkten Potenzreihe", Göttinger Nachrichten (1916), S. 276—291 and (1917), S. 119—128. See also a paper by L. Neder, with the same title, Math. Zeitschr. **11** (1921), S. 115—123.

[3]) $f^1(z)$ is not quite the ordinary derivative $f'(z)$, just as $f^{\alpha}(z)$ is not quite the Liouville or Hadamard derivative of fractional order; but the differences are, for our purposes, trivial, and we ignore them. See § 2. 1 of P.

[4]) This follows (if we pay attention to the glosses in § 2. 1 of P) from Theorem 33 of our paper **21** (in the bibliography of P), with $1 - \alpha$ for α, $\frac{1}{2}$ for p, and so $\frac{1}{1+\alpha}$ for $q = \frac{p}{1 - p\alpha}$.

We define $g(z)$ and $f_N(z)$ as in § 2, but now take

$$f(z) = \sum_1^\infty \frac{1}{r^s} f_{N_r}(z), \tag{6}$$

where (N_r) is an increasing sequence of numbers which satisfy (2) and tend to infinity so rapidly that

$$\frac{\log N_r}{r^s(N_1 + N_2 + \ldots + N_{r-1})} \to \infty \tag{7}$$

and *a fortiori*

$$\frac{N_r}{N_{r-1}^2} \to \infty. \tag{8}$$

Then, first,

$$\int |f^1|^{\frac{1}{2}} d\theta \leqq \sum_1^\infty r^{-\frac{1}{2}} \int |f_{N_r}^1|^{\frac{1}{2}} d\theta < K. \tag{9}$$

If we write

$$f_{N_r}(z) = \sum a_{n,r} z^n$$

(so that $a_{n,r} \geqq 0$ and $a_{n,r} = 0$ if $n > 3N_r + 1$), we have

$$\sum a_{n,r} = \frac{1}{N_r + 1} \int_0^1 \left(\frac{1 - z^{N_r+1}}{1-z} \right)^3 dz < K N_r, \tag{10}$$

and

$$\sum_1^{N_r} a_{n,s} = \frac{1}{2(N_s+1)} \sum_1^{N_r} (n+1) < K \frac{N_r^2}{N_s} \tag{11}$$

if $s > r$. Also $|b_n| \leqq 1$. Hence, if we write

$$\sum_1^{N_r} a_n b_n = \sum_1^{N_r} b_n \sum_{s=1}^\infty \frac{a_{n,s}}{s^s} = T_1 + T_2 + T_3, \tag{12}$$

where T_1, T_2, and T_3 are the contributions of the terms for which $s < r$, $s = r$ and $s > r$ respectively, we have

$$|T_1| \leqq \sum_{s=1}^r \frac{1}{s^s} \sum a_{n,s} < K(N_1 + N_2 + \ldots + N_{r-1}), \tag{13}$$

by (10), and

$$|T_3| \leqq \sum_{s=r+1}^\infty \frac{1}{s^s} \sum_{n=1}^{N_r} a_{n,s} < K \frac{N_r^2}{N_{r+1}}, \tag{14}$$

by (11). It follows from (14) and (8) that $T_3 \to 0$. Also, by (5),

$$|T_2| > K \frac{\log N_r}{r^r}; \tag{15}$$

and (13), (15), and (7) show that $\sum a_n b_n$ is divergent.

(Eingegangen am 5. September 1931.)

On a Problem in the additive Theory of Numbers.

(Third paper.)

By

C. J. A. Evelyn in Wotton and E. H. Linfoot in Princeton.

---

1. Introductory. In two recent papers[1]) we considered the representation of a large number as the sum of the M-numbers defined as follows:

Definition. The integer

$$n = p_1^{\lambda_1} p_2^{\lambda_2} \dots p_\varrho^{\lambda_\varrho}$$

is said to be a "number of class N" or "M-number" when $\lambda_1, \lambda_2, \dots, \lambda_\varrho \leqq N-1$.

In the first paper we showed by the analytic method of Hardy and Littlewood that for $s \geqq 3$ the number of representations of a large number n as the sum of s M-numbers

$$r_s(n) = \frac{n^{s-1}}{(s-1)!\,\zeta^s(N)} S_s(n) + O\left(n^{s-\frac{3}{2}+\frac{1}{2N}+\varepsilon}\right) \tag{1.11}$$

as $n \to \infty$, for every $\varepsilon > 0$, where

$$S_s(n) = \prod_{p^N+n}\left(1+\frac{(-1)^{s+1}}{(p^N-1)^s}\right)\prod_{p^N \mid n}\left(1+\frac{(-1)^s}{(p^N-1)^{s-1}}\right). \tag{1.12}$$

In the second paper we discussed the case $s=2$ by a purely elementary argument, and showed that

$$r_2(n) = \frac{n}{\zeta^2(N)} S_2(n) + O\left(n^{\frac{2}{N+1}+\varepsilon}\right) \tag{1.2}$$

as $n \to \infty$, for every $\varepsilon > 0$, where $S_2(n)$ is given by (1.12). A simpler proof of this result has been given by T. Estermann[2]). In the present paper we show how, starting from (1.2), we can derive by strictly elementary arguments an asymptotic formula for $r_s(n)$ for every $s \geqq 3$ and prove

---

[1]) Math. Zeitschr. **30** (1929), S. 433–448; Journ. f. Math. **164** (1931), S. 131–140.

[2]) Journal London Math. Soc. **6** (1931), p. 37–40.

Theorem 1. *For* $N \geqq 2$, $s \geqq 2$

$$(1.3) \qquad r_s(n) = \frac{n^{s-1}}{(s-1)!}\frac{1}{\zeta^s(N)} S_s(n) + O\left(n^{s-2+\frac{2}{N+1}+\varepsilon}\right),$$

as $n \to \infty$, *for every* $\varepsilon > 0$.

2. We first prove two lemmas.

Lemma 2.1. *If* $s \geqq 0$ *is an integer,* $0 < \theta < 1$ *and* $T > 0$, *then*

$$\left| \sum_{0 \leqq m+\theta \leqq T} (m+\theta)^s - \frac{T^{s+1}}{s+1} \right| < c_1 T^s,$$

where summation is over the integer m *and* $c_1 > 0$ *depends on* s *only,* (as do $c_2, \ldots, c_8 > 0$).

Proof. 1. We first dispose of the case $0 < T < 1$. There are two possibilities, $0 < T < \theta$ and $\theta \leqq T < 1$. When the first holds we have to prove $\frac{T^{s+1}}{s+1} < c_2 T^s$, which is trivial with $c_2 = 1$. When the second holds we have to prove

$$\left| \theta^s - \frac{T^{s+1}}{s+1} \right| < c_3 T^s,$$

i. e.

$$(s+1)\left(\frac{\theta}{T}\right)^s - T$$

bounded in $0 < \theta \leqq T < 1$, which again is trivial. Thus we may suppose $T \geqq 1$.

2. The proof now proceeds by induction. For $s = 0$ the result is immediate. Let $s \geqq 1$ and suppose the lemma proved for $s-1$. Thus for $T \geqq 1$

$$\left| \sum_{0 \leqq m+\theta \leqq T} (m+\theta)^{s-1} - \frac{T^s}{s} \right| < c_4 T^{s-1},$$

from which it follows that for $m \geqq 1$

$$\left| \sum_{0 \leqq \nu+\theta \leqq m} (\nu+\theta)^{s-1} - \frac{(m+\theta)^s}{s} \right| < c_5 m^{s-1},$$

the summation being over ν. Writing

$$S(m,\theta) = \sum_{0 \leqq \nu+\theta \leqq m} (\nu+\theta)^{s-1},$$

so that $S(0,\theta)$ and $S(-1,\theta)$ are both zero, we have

$$\begin{aligned} \sum_{0 \leqq m+\theta \leqq T} (m+\theta)^s &= \sum_{m=0}^{T-\theta} (m+\theta)\,(S(m,\theta) - S(m-1,\theta)) \\ &= -\sum_{m=0}^{T-\theta} S(m,\theta) + ([T-\theta]+1)\,S(T-\theta,\theta). \end{aligned}$$

Here

$$\left| \sum_{m=0}^{T-\theta} \left(S(m,\theta) - \frac{(m+\theta)^s}{s} \right) \right| < c_5 \sum_{m=0}^{T-\theta} m^{s-1} < c_6 T^s$$

and

$$\left| ([T-\theta]+1)\, S(T-\theta,\theta) - \frac{T^{s+1}}{s} \right| < c_7 T^s.$$

Thus

$$\left| \sum_{0 \leqq m+\theta \leqq T} (m+\theta)^s + \sum_{m=0}^{T-\theta} \frac{(m+\theta)^s}{s} - \frac{T^{s+1}}{s} \right| < c_6 T^s + c_7 T^s,$$

$$\left| \frac{s+1}{s} \sum_{0 \leqq m+\theta \leqq T} (m+\theta)^s - \frac{T^{s+1}}{s} \right| < (c_6 + c_7) T^s,$$

$$\left| \sum_{0 \leqq m+\theta \leqq T} (m+\theta)^s - \frac{T^{s+1}}{s+1} \right| < c_8 T^s$$

and the lemma is proved.

Lemma 2.2. *For* $s \geqq 2$

$$(2.21) \qquad \sum_{(m_1,\ldots,m_s)^N \mid n} \frac{\mu(m_1)\ldots\mu(m_s)}{m_1^N \ldots m_s^N} (m_1,\ldots,m_s)^N = \frac{1}{\zeta^s(N)} S_s(n).$$

Proof. Let $\varrho = \prod_{p^N \mid n} p$. Then $(m_1,\ldots,m_s)^N \mid n$ is equivalent to $(m_1,\ldots,m_s) \mid \varrho$ and the left hand side of (2.21) is

$$\sum_{m_1=1}^{\infty} \frac{\mu(m_1)}{m_1^N} \cdots \sum_{m_{s-1}=1}^{\infty} \frac{\mu(m_{s-1})}{m_{s-1}^N} \sum_{\substack{m_s=1 \\ (m_1,\ldots,m_s) \mid \varrho}}^{\infty} \frac{\mu(m_s)}{m_s^N} (m_1,\ldots,m_s)^N.$$

The inmost sum can be written as

$$\prod_{\substack{p \\ (m_1,\ldots,m_{s-1},p) \mid \varrho}} \left(1 - \frac{(m_1,\ldots,m_{s-1},p)^N}{p^N} \right),$$

since its summand is multiplicative,

$$= \prod_{p \nmid (m_1,\ldots,m_{s-1})} \left(1 - \frac{1}{p^N}\right) \prod_{\substack{p \mid (m_1,\ldots,m_{s-1}) \\ p \mid \varrho}} (1-1)$$

$$\begin{cases} = 0 & \text{if } (m_1,\ldots,m_{s-1},\varrho) > 1, \\ = \prod_{p \nmid (m_1,\ldots,m_{s-1})} \left(1 - \frac{1}{p^N}\right) & \text{if } (m,\ldots,m_{s-1},\varrho) = 1. \end{cases}$$

Thus our expression is

$$\frac{1}{\zeta(N)} \sum_{m_1=1}^{\infty} \frac{\mu(m_1)}{m_1^N} \cdot \ldots \cdot \sum_{\substack{m_{s-1}=1 \\ (m_1,\ldots,m_{s-1},\varrho)=1}}^{\infty} \frac{\mu(m_{s-1})}{m_{s-1}^N} \prod_{p \mid (m_1,\ldots,m_{s-1})} \frac{p^N}{p^N - 1}$$

The inmost sum again has a multiplicative summand and so

$$= \prod_{\substack{p \\ (m_1, \ldots, m_{s-2}, p, \varrho)=1}} \left(1 - \frac{1}{p^N} \prod_{\pi \mid (m_2, \ldots, m_{s-1}, p)} \frac{\pi^N}{\pi^N - 1}\right)$$

$$= \prod_{\substack{p \mid (m_1, \ldots, m_{s-2}) \\ p \nmid \varrho}} \left(1 - \frac{1}{p^N - 1}\right) \prod_{p \nmid (m_1, \ldots, m_{s-2})} \left(1 - \frac{1}{p^N}\right)$$

$$= \frac{1}{\zeta(N)} \prod_{\substack{p \mid (m_1, \ldots, m_{s-2}) \\ p \nmid \varrho}} \frac{p^N - 2}{p^N - 1} \prod_{p \mid (m_1, \ldots, m_{s-2})} \frac{p^N}{p^N - 1}$$

$$= \frac{1}{\zeta(N)} \prod_{\substack{p \mid (m_1, \ldots, m_{s-2}) \\ p \nmid \varrho}} \frac{p^{2N} - 2p^N}{(p^N - 1)^2} \prod_{\substack{p \mid (m_1, \ldots, m_{s-2}) \\ p \mid \varrho}} \frac{p^N}{p^N - 1}$$

$$= \frac{1}{\zeta(N)} \prod_{\substack{p \mid (m_1, \ldots, m_{s-2}) \\ p \nmid \varrho}} \left(1 - \frac{1}{(p^N - 1)^2}\right) \prod_{\substack{p \mid (m_1, \ldots, m_{s-2}) \\ p \mid \varrho}} \left(1 + \frac{1}{p^N - 1}\right)$$

and our expression becomes

$$\frac{1}{\zeta^2(N)} \sum_{m_1=1}^{\infty} \frac{\mu(m_1)}{m_1^N} \cdots \sum_{m_{s-2}=1}^{\infty} \frac{\mu(m_{s-2})}{m_{s-2}^N} \prod_{\substack{p \mid (m_1, \ldots, m_{s-2}) \\ p \nmid \varrho}} \left(1 - \frac{1}{(p^N - 1)^2}\right) \prod_{\substack{p \mid (m_1, \ldots, m_{s-2}) \\ p \mid \varrho}} \left(1 + \frac{1}{p^N}\right.$$

When $s = 2$ the Σ's are absent and we have $\frac{1}{\zeta^2(N)} S_2(n)$. If $s > 2$ and $2 \leqq r < s$, then

$$\sum_{m_{s-r}=1}^{\infty} \frac{\mu(m_{s-r})}{m_{s-r}^N} \prod_{\substack{p \mid (m_1, \ldots, m_{s-r}) \\ p \nmid \varrho}} \left(1 + \frac{(-1)^{r+1}}{(p^N - 1)^r}\right) \prod_{\substack{p \mid (m_1, \ldots, m_{s-r}) \\ p \mid \varrho}} \left(1 + \frac{(-1)^r}{(p^N - 1)^{r-1}}\right)$$

has a multiplicative summand and can be written

$$\prod_{p \nmid \varrho} \left\{1 - \frac{1}{p^N} \prod_{\substack{\pi \mid (m_1, \ldots, m_{s-r-1}, p) \\ \pi \nmid \varrho}} \left(1 + \frac{(-1)^{r+1}}{(\pi^N - 1)^r}\right)\right\}$$

$$\prod_{p \mid \varrho} \left\{1 - \frac{1}{p^N} \prod_{\substack{\pi \mid (m_1, \ldots, m_{s-r-1}, p) \\ \pi \mid \varrho}} \left(1 + \frac{(-1)^r}{(\pi^N - 1)^{r-1}}\right)\right\}$$

$$= \prod_{\substack{p \mid (m_1, \ldots, m_{s-r-1}) \\ p \nmid \varrho}} \left\{1 - \frac{1}{p^N}\left(1 + \frac{(-1)^{r+1}}{(p^N - 1)^r}\right)\right\} \prod_{p \nmid (m_1, \ldots, m_{s-r-1})} \left(1 - \frac{1}{p^N}\right)$$

$$\prod_{\substack{p \mid (m_1, \ldots, m_{s-r-1}) \\ p \mid \varrho}} \left\{1 - \frac{1}{p^N}\left(1 + \frac{(-1)^r}{(p^N - 1)^{r-1}}\right)\right\} \prod_{\substack{p \nmid (m_1, \ldots, m_{s-r-1}) \\ p \mid \varrho}} \left(1 - \frac{1}{p^N}\right)$$

$$= \prod_{p \nmid (m_1, \ldots, m_{s-r-1})} \left(1 - \frac{1}{p^N}\right) \prod_{\substack{p \mid (m_1, \ldots, m_{s-r-1}) \\ p \nmid q}} \left\{\frac{p^N - 1}{p^N} + \frac{(-1)^{r+2}}{p^N (p^N - 1)^r}\right\}$$

$$\prod_{\substack{p \mid (m_1, \ldots, m_{s-r-1}) \\ p \mid q}} \left\{\frac{p^N - 1}{p^N} + \frac{(-1)^{r+1}}{p^N (p^N - 1)^{r-1}}\right\}$$

$$= \frac{1}{\zeta(N)} \prod_{\substack{p \mid (m_1, \ldots, m_{s-r-1}) \\ p \nmid q}} \left(1 + \frac{(-1)^{r+2}}{(p^N - 1)^{r+1}}\right) \prod_{\substack{p \mid (m_1, \ldots, m_{s-r-1}) \\ p \mid q}} \left(1 + \frac{(-1)^{r+1}}{(p^N - 1)^r}\right).$$

Applying this result successively for $r = 2, 3, \ldots, s-1$ our expression becomes

$$\frac{1}{\zeta^s(N)} \prod_{p \nmid q} \left(1 + \frac{(-1)^{s+1}}{(p^N - 1)^s}\right) \prod_{p \mid q} \left(1 + \frac{(-1)^s}{(p^N - 1)^{s-1}}\right)$$

$$= \frac{1}{\zeta^s(N)} \prod_{p^N \nmid n} \left(1 + \frac{(-1)^{s+1}}{(p^N - 1)^s}\right) \prod_{p^N \mid n} \left(1 + \frac{(-1)^s}{(p^N - 1)^{s-1}}\right)$$

and the lemma is proved.

3. Proof of Theorem 1.

1. When $s = 2$ the result follows by (1.2).

2. Let $s \geqq 3$ and suppose the theorem proved for $\nu_{s-1}(n)$. Thus

$$\nu_{s-1}(n) = \frac{n^{s-2}}{(s-2)!} \frac{1}{\zeta^{s-1}(N)} S_{s-1}(n) + O\left(n^{s-3+\frac{3}{N+1}+\varepsilon}\right)$$

for every $\varepsilon > 0$. And

$$(n) = \sum_{M_1 + M_2 + \ldots + M_s = n} 1 = \sum_{M_s=1}^{n} \nu_{s-1}(n - M_s)$$

$$= \sum_{M_s=1}^{n} \frac{(n - M_s)^{s-2}}{(s-2)!} \frac{1}{\zeta^{s-1}(N)} S_{s-1}(n - M_s) + O \sum_{M_s=1}^{n} (n - M_s)^{s-3+\frac{2}{N+1}+\varepsilon}$$

$$= \frac{1}{(s-2)!} \sum_{M_s=1}^{n} (n - M_s)^{s-2} \sum_{(m_1, \ldots, m_{s-1})^N \mid n - M_s} \frac{\mu(m_1) \ldots \mu(m_{s-1})}{m_1^N \ldots m_{s-1}^N} (m_1, \ldots, m_{s-1})^N,$$

by (2.2). Since

$$\sum_{m^N \mid r} \mu(m) = 1 \quad r \text{ an } M\text{-number}$$
$$= 0. \quad \text{otherwise,}$$

the first term is

$$\frac{1}{(s-2)!}\sum_{r=1}^{n}(n-r)^{s-2}\sum_{m_s^N\mid r}\mu(m_s)\sum_{(m_1,\ldots,m_{s-1})^N\mid n-r}\frac{\mu(m_1)\ldots\mu(m_{s-1})}{m_1^N\ldots m_{s-1}^N}(m_1,\ldots,m_{s-1})^N$$

$$=\frac{1}{(s-2)!}\sum_{m_s=1}^{n^{1/N}}\mu(m_s)\sum_{r\equiv 0 \bmod m_s^N}^{n}(n-r)^{s-2}\sum_{(m_1,\ldots,m_{s-1})^N\mid n-r}\frac{\mu(m_1)\ldots\mu(m_{s-1})}{m_1^N\ldots m_{s-1}^N}(m_1,\ldots,$$

$$=\frac{1}{(s-2)!}\sum_{m_s=1}^{n^{1/N}}\mu(m_s)\sum_{\nu=1}^{n/m_s^N}(n-m_s^N\nu)^{s-2}\sum_{(m_1,\ldots,m_{s-1})^N\mid n-m_s^N\nu}\frac{\mu(m_1)\ldots\mu(m_{s-1})}{m_1^N\ldots m_{s-1}^N}(m_1,\ldots,$$

$$=\frac{1}{(s-2)!}\sum_{m_s=1}^{n^{1/N}}\mu(m_s)\sum_{(m_1,\ldots,m_{s-1})^N\leqq n}\frac{\mu(m_1)\ldots\mu(m_{s-1})}{m_1^N\ldots m_{s-1}^N}(m_1,\ldots,m_{s-1})^N$$

$$\times\sum_{\substack{\nu=1\\ m_s^N\nu\equiv n \bmod (m_1,\ldots,m_{s-1})^N}}^{n/m_s^N}(n-m$$

Now

$$m_s^N\nu\equiv n \mod (m_1,\ldots,m_{s-1})^N$$

has no solution in ν unless $(m_1,\ldots,m_{s-1},m_s)^N\mid n$; when this is satisfied the congruence can be written

$$\frac{m_s^N}{(m_1,\ldots,m_s)^N}\nu\equiv\frac{n}{(m_1,\ldots,m_s)^N} \mod \frac{(m_1,\ldots,m_{s-1})^N}{(m_1,\ldots,m_s)^N}$$

with $\dfrac{m_s^N}{(m_1,\ldots,m_s)^N}$ prime to $\dfrac{(m_1,\ldots,m_{s-1})^N}{(m_1,\ldots,m_s)^N}$; this in turn is equivalent to

$$\nu\equiv\frac{n}{(m_1,\ldots,m_s)^N}L \mod \frac{(m_1,\ldots,m_{s-1})^N}{(m_1,\ldots,m_s)^N},$$

where L is the socius $\mod\dfrac{(m_1,\ldots,m_{s-1})^N}{(m_1,\ldots,m_s)^N}$ of $\dfrac{m_s^N}{(m_1,\ldots,m_s)^N}$; that is to say $0<L<\dfrac{(m_1,\ldots,m_{s-1})^N}{(m_1,\ldots,m_s)^N}$ and

$$\frac{m_s^N L}{(m_1,\ldots,m_s)^N}\equiv 1 \mod \frac{(m_1,\ldots,m_{s-1})^N}{(m_1,\ldots,m_s)^N}.$$

Writing

$$n-m_s^N\nu=\frac{(m_1,\ldots,m_{s-1})^N}{(m_1,\ldots,m_s)^N}m_s^N(t+\theta),$$

where t is an integer and $0 < \theta = \frac{(m_1, \ldots, m_s)^N L}{(m_1, \ldots, m_{s-1})^N} < 1$, our sum becomes

$$\frac{1}{(s-2)!} \sum_{m_s=1}^{n^{1/N}} \mu(m_s) \sum_{\substack{(m_1, \ldots, m_{s-1})^N \leqq n \\ (m_1, \ldots, m_s)^N \mid n}} \frac{\mu(m_1) \ldots \mu(m_{s-1})}{m_1^N \ldots m_{s-1}^N} (m_1, \ldots, m_{s-1})^N$$

$$\sum_{0 \leqq n - m_s^N \nu \leqq n-1} \frac{(m_1, \ldots, m_{s-1})^{(s-2)N}}{(m_1, \ldots, m_s)^{(s-2)N}} m_s^{(s-2)N} (t+\theta)^{s-2}$$

$$= \frac{1}{(s-2)!} \sum_{\substack{(m_1, \ldots, m_{s-1})^N \leqq n,\ m_s^N \leqq n \\ (m_1, \ldots, m_s)^N \mid n}} \frac{\mu(m_1) \ldots \mu(m_s)}{m_1^N \ldots m_s^N} \frac{(m_1, \ldots, m_{s-1})^{(s-1)N}}{(m_1, \ldots, m_s)^{(s-2)N}} m_s^{(s-1)N}$$

$$\sum_{0 \leqq \frac{(m_1, \ldots, m_{s-1})^N}{(m_1, \ldots, m_s)^N} m_s^N (t+\theta) \leqq n-1} (t+\theta)^{s-2}.$$

The inner sum is

$$\frac{n^{s-1}}{s-1} \frac{(m_1, \ldots, m_s)^{(s-1)N}}{(m_1, \ldots, m_{s-1})^{(s-1)N}} \frac{1}{m_s^{(s-1)N}} + O\left(\frac{n^{s-2} (m_1, \ldots, m_s)^{(s-2)N}}{(m_1, \ldots, m_{s-1})^{(s-2)N} m_s^{(s-2)N}} \right)$$

by lemma 2.1. Thus our expression is

$$\frac{n^{s-1}}{(s-1)!} \sum_{\substack{(m_1, \ldots, m_{s-1})^N \leqq n,\ m_s^N \leqq n \\ (m_1, \ldots, m_s)^N \mid n}} \frac{\mu(m_1) \ldots \mu(m_s)}{m_1^N \ldots m_s^N} (m_1, \ldots, m_s)^N$$

$$+ O\left(n^{s-2} \sum_{\substack{(m_1, \ldots, m_{s-1})^N \leqq n,\ m_s^N \leqq n \\ (m_1, \ldots, m_s)^N \mid n}} \frac{(m_1, \ldots, m_{s-1})^N}{m_1^N \ldots m_{s-1}^N} \right).$$

Now

$$\sum_{\substack{(m_1, \ldots, m_{s-1})^N \leqq n,\ m_s^N \leqq n \\ (m_1, \ldots, m_s)^N \mid n}} \frac{(m_1, \ldots, m_{s-1})^N}{m_1^N \ldots m_{s-1}^N} < n^{1/N} \sum_{m_1, \ldots, m_{s-1} \geqq 1} \frac{(m_1, \ldots, m_{s-1})^N}{m_1^N \ldots m_{s-1}^N}$$

$$= n^{1/N} \sum_{d=1}^{\infty} \sum_{(m_1, \ldots, m_{s-1}) = d} \frac{(m_1, \ldots, m_{s-1})^N}{m_1^N \ldots m_{s-1}^N} = n^{1/N} \sum_{d=1}^{\infty} \frac{1}{d^{(s-2)N}} \sum_{(r_1, \ldots, r_{s-1}) = 1} \frac{1}{r_1^N \ldots r_{s-1}^N}$$

$$= O(n^{1/N}),$$

since $s \geqq 3$, while

$$\left| \left(\sum_{(m_1, \ldots, m_s)^N \mid n} - \sum_{\substack{(m_1, \ldots, m_{s-1})^N \leqq n,\ m_s^N \leqq n \\ (m_1, \ldots, m_s)^N \mid n}} \right) \frac{\mu(m_1) \ldots \mu(m_s)}{m_1^N \ldots m_s^N} (m_1, \ldots, m_s)^N \right|$$

$$< \sum_{\substack{m_s^N > n \\ m_1, \ldots, m_{s-1} \geqq 1}} \frac{(m_1, \ldots, m_s)^N}{m_1^N \ldots m_s^N} + \sum_{\substack{(m_1, \ldots, m_{s-1})^N > n \\ m_s \geqq 1}} \frac{(m_1, \ldots, m_s)^N}{m_1^N \ldots m_s^N}.$$

The first term

$$< \sum_{m_s^N > n} \frac{1}{m_s^N} \sum_{m_1, \ldots, m_{s-1} \geqq 1} \frac{(m_1, \ldots, m_{s-1})^N}{m_1^N \ldots m_{s-1}^N}$$

$$= O\left(n^{-1+\frac{1}{N}}\right),$$

since $s \geqq 3$; the second

$$= \sum_{d=1}^{\infty} \sum_{\substack{(m_1, \ldots, m_s) = d \\ (m_1, \ldots, m_{s-1})^N > n}} \frac{(m_1, \ldots, m_s)^N}{m_1^N \ldots m_s^N}$$

$$= \sum_{d=1}^{\infty} \frac{1}{d^{N(s-1)}} \sum_{\substack{(r_1, \ldots, r_s) = 1 \\ (r_1, \ldots, r_{s-1})^N > n/d^N}} \frac{1}{r_1^N \ldots r_s^N}$$

$$< \sum_{d=1}^{\infty} \frac{1}{d^{N(s-1)}} \sum_{r_s=1}^{\infty} \frac{1}{r_s^N} \sum_{r_{s-1}^N > n\, d^N} \frac{1}{r_{s-1}^N} \sum_{r_1, \ldots, r_{s-2} \geqq 1} \frac{1}{r_1^N \ldots r_{s-2}^N}$$

$$= \zeta^{s-1}(N) \sum_{d=1}^{\infty} \frac{1}{d^{N(s-1)}} O\left(n^{-1+\frac{1}{N}} d^{N-1}\right)$$

$$= O\left(n^{-1+\frac{1}{N}} \sum_{d=1}^{\infty} \frac{1}{d^{N(s-2)+1}}\right)$$

$$= O\left(n^{-1+\frac{1}{N}}\right).$$

Thus our sum is

$$\frac{n^{s-1}}{(s-1)!} \sum_{(m_1, \ldots, m_s)^N \mid n} \frac{\mu(m_1) \ldots \mu(m_s)}{m_1^N \ldots m_s^N} (m_1, \ldots, m_s)^N + O\left(n^{s-2+\frac{1}{N}}\right)$$

$$= \frac{n^{s-1}}{(s-1)!} \frac{1}{\zeta^s(N)} S_s(n) + O\left(n^{s-2+\frac{1}{N}}\right)$$

by (2.2), and hence from (3.1)

$$r_s(n) = \frac{n^{s-1}}{(s-1)!} \frac{1}{\zeta^s(N)} S_s(n) + O\left(n^{s-2+\frac{1}{N+1}+\varepsilon}\right).$$

The theorem is therefore proved by induction.

(Eingegangen am 1. September 1930.)

Minimaxeigenschaft der offenen und geschlossenen Extremalen.

Von

Michael Kerner in Warschau.

1. Einleitung.

In der Arbeit „Dynamical Systems with two Degrees of Freedom“[1] hat Herr Birkhoff den Begriff der geschlossenen Bahnen vom Typus des „Minimax“ eingeführt. Sie haben in manchen Fällen zu Existenzbeweisen gedient. Diese Minimaxeigenschaft läßt sich leicht auf das Problem der Variationsrechnung übertragen, das die Extremalen, die zwei Punkte verbinden, wie auch die geschlossenen Extremalen betrifft.

Die vorliegende Arbeit ist diesen beiden Extremalarten in der Ebene gewidmet. Doch sind nicht die Existenzsätze ihr Gegenstand, sondern die notwendigen und hinreichenden Bedingungen der Minimaxeigenschaft einer Extremale, die entweder zwei Punkte verbindet oder geschlossen ist. Die hinreichenden Bedingungen sind der Hauptgegenstand der Arbeit. Die notwendigen sind nur der Vollständigkeit halber angeführt.

Was die Beweise der notwendigen Bedingungen in beiden Fällen betrifft, so unterscheiden sie sich nicht wesentlich von den von Herrn Birkhoff im Falle der geschlossenen Bahnen gegebenen.

Die Minimaxeigenschaft selbst wird in der vorliegenden Arbeit etwas anders definiert. Bekanntlich zählt Herr Birkhoff eine geschlossene Bahn zum Minimaxtypus, wenn man in einer hinreichend kleinen Umgebung zwei Bahnen finden kann, längs deren das definite Integral kleinere Werte annimmt als längs der gegebenen Bahn und die sich nicht stetig in der Umgebung so ineinander überführen lassen, daß das Integral auch in allen Übergangsstadien kleinere Werte annehme als längs der gegebenen Bahn. Wir sagen dagegen, daß eine gegebene Extremale einer bestimmten Klasse

[1]) Transactions of the American Mathematical Society 18 (1917), S. 239 u. f.

die Minimaxeigenschaft besitzt, wenn es in einer hinreichend kleinen Umgebung zwei Kurven derselben Klasse gibt, die *auf beiden Seiten* der gegebenen Extremale liegen und die im übrigen alle Forderungen der vorigen Definition erfüllen. Wir werden selbstverständlich in beiden Fällen genau definieren, was man unter den Worten „auf beiden Seiten liegen" verstehen soll.

Wir übergehen in der Arbeit gänzlich die Beweise dafür, daß jede Kurve einer Klasse sich in jede andere überführen läßt.

Erstes Kapitel.

Die Extremalen, die zwei Punkte verbinden.

2. Voraussetzungen und Definitionen.

Betrachten wir das Integral

$$J = \int F\left(x, y, \frac{dx}{dt}, \frac{dy}{dt}\right) dt, \tag{1}$$

wo die Funktion $F(x, y, p, q)$ in einem gewissen Bereiche R[2]) der x, y-Ebene P für alle nichtzugleichverschwindenden Werte von p, q von der Klasse C'' und positiv-homogen vom ersten Grade ist.

Im Bereiche R befinde sich eine nichtgebrochene Extremale $\mathfrak{E}$ ohne mehrfache Punkte, die zwei Punkte A und B verbindet. Der Einfachheit halber stellen wir $\mathfrak{E}$ durch Gleichungen

$$x = \varphi(t), \quad y = \psi(t), \quad 0 \leqq t \leqq d \tag{2}$$

dar, wo t die Bogenlänge vom Punkte A ab bezeichnet.

Wir setzen voraus, daß längs $\mathfrak{E}$

$$F(\varphi, \psi, p, q) > 0 \tag{3}$$

für alle nichtzugleichverschwindenden p, q, daß

$$F_1(\varphi, \psi, \varphi_t, \psi_t) > 0 \tag{4}$$

und daß

$$\mathcal{E}(\varphi, \psi, \varphi_t, \psi_t, p, q) > 0 \tag{5}$$

für alle nicht zu φ_t, ψ_t proportionalen p, q.

Bekanntlich sind dann $\varphi(t), \psi(t)$ von der Klasse C''. Wir verkleinern den Bereich R so, daß (3) im ganzen R erfüllt ist.

Als eine Kurve werden wir im folgenden nur eine streckbare Kurve verstehen. Wir können sie in der Form darstellen

$$x = x(\tau), \quad y = y(\tau), \quad a \leqq \tau \leqq b, \tag{6}$$

[2]) Unter einem Bereiche werden wir immer eine offene, zusammenhängende Punktmenge verstehen.

o $x(\tau), y(\tau)$ totalstetig sind. Sie kann auch mehrfache Punkte enthalten. allen die Punkte, die den Werten a und b entsprechen, zusammen, so wird ie Kurve geschlossen genannt. Ist $\mathfrak{L}$ eine Kurve, so werden wir mit $\overleftarrow{\mathfrak{L}}$ ieselbe im entgegengesetzten Sinne durchlaufen bezeichnen. Bezeichnet das ymbol $\int$ das Lebesguesche Integral, so hat (1) für jede Kurve einen Sinn nd ist gegen Parametervertauschung invariant.

Sei M eine beliebige Punktmenge. Mit $U(M)$ werden wir die Umebung von M bezeichnen, mit $U_\varrho(M)$ speziell die Menge der Punkte, deren leinste Entfernung von M kleiner ist als ϱ. Wenn es sich um die Umgebung er ganzen Extremalen $\mathfrak{E}$ handelt, werden wir U, U_ϱ ohne Klammern schreiben.

Entfernt man von der Ebene P alle Punkte eines abgeschlossenen eilbogens $\mathfrak{C}$ von $\mathfrak{E}$, so bekommt man einen zweifach zusammenhängenden ereich $(P - \mathfrak{C})$.

Definition 1. *Wir sagen, daß zwei Kurven $\mathfrak{L}_1$ und $\mathfrak{L}_2$, welche die unkte A und B verbinden, auf beiden Seiten der Kurve $\mathfrak{E}$ liegen, wenn ie geschlossene Kurve, die aus $\overleftarrow{\mathfrak{L}_1}$ und $\mathfrak{L}_2$ zusammengesetzt ist, irreduzibel nicht zusammenziehbar) in bezug auf jeden Bereich $(P - \mathfrak{C})$ ist, wo $\mathfrak{C}$ jeden, eder A noch B enthaltenden, abgeschlossenen Teilbogen von $\mathfrak{E}$ bezeichnet.*

Man könnte diese Definition vereinfachen, indem man $\mathfrak{E}$ anstatt $\mathfrak{C}$ ähme. Doch würde dann die geschlossene Kurve nicht ganz in $(P - \mathfrak{E})$ egen. Ich glaube, daß die Definition 1 möglichst allgemein ist. Sie entheidet nicht, auf welcher Seite von $\mathfrak{E}$ ein Punkt liegt.

Definition 2. *Wir sagen, daß zwei Kurven einer Klasse sich in dieser lasse stetig ineinander überführen lassen, wenn eine Schar von Kurven ieser Klasse existiert, die von einem Parameter in einem gewissen Interalle stetig (und zwar gleichmäßig in bezug auf τ) abhängt und wenn die egebenen Kurven zwei Werten dieses Intervalles entsprechen.*

Definition 3. *Wir sagen, daß die Extremale $\mathfrak{E}$ die Minimaxeigenschaft esitzt, wenn in einer hinreichend kleinen Umgebung U von $\mathfrak{E}$ die folgende edingung erfüllt ist:*

Wenn es zwei Kurven gibt, die A und B verbinden, die in U auf eiden Seiten von $\mathfrak{E}$ liegen und längs deren das Integral J einen kleineren ert annimmt als längs $\mathfrak{E}$, so lassen sich diese Kurven nicht ineinander etig in der Klasse der Kurven überführen, die A und B verbinden, in U egen und längs deren das Integral J kleiner bleibt als längs $\mathfrak{E}$.

Es ist leicht einzusehen, daß unsere Definition der Minimaxeigenschaft ıch die Fälle des Minimums und des einseitigen Minimums umfaßt, weil in diesen Fällen keine zwei Kurven gibt, von denen in der Vorbedingung e Rede ist. In dieser Hinsicht unterscheidet sich unsere Definition von er von Herrn Birkhoff.

Noch einige Worte über die Schreibweise.

Wir werden den Integrationsweg bei J in Klammern hinzufügen, w etwa $J(\mathfrak{E})$, $J(A\mathfrak{L}_1 B)$ usw.

Ist P ein Punkt der Extremale $\mathfrak{E}$ vom Parameterwert t, so bezeichne $P^{(1)}, P^{(2)}, P^{(3)}, \ldots$ und $t^{(1)}, t^{(2)}, t^{(3)}, \ldots$ die nachfolgenden konjugierten Punkt nebst ihren Parameterwerten. Entsprechend bezeichnen $P^{(-1)}, P^{(-2)}, P^{(-3)}, .$ und $t^{(-1)}, t^{(-2)}, t^{(-3)}, \ldots$ die vorangehenden konjugierten Punkte und ih Parameterwerte.

3. Hilfssatz.

Auf der Extremale $A\mathfrak{E}B$ (Fig. 1) wählen wir zwei Punkte C und so, daß D dem ersten konjugierten Punkte $C^{(1)}$ von C vorangeht (wen überhaupt $C^{(1)}$ existiert). Bekanntlich kann man unserer Voraussetzung halber eine Umgebung des Bogens $C\mathfrak{E}D$ finden, in der der Bogen $C\mathfrak{E}$ ein Minimum des Integrals J liefert.

Ich behaupte, daß es eine Umgebung des ganzen Bogens $A\mathfrak{E}B$ gib in der der Bogen $C\mathfrak{E}D$ ein Minimum des Integrals J liefert.

Beweis. Da D dem (eventuell existierenden) Punkte $C^{(1)}$ vorangeh kann man den Bogen $C\mathfrak{E}D$ in eine reguläre Extremalenschar einbette Führen wir zu dieser Schar die Transversalen $\mathfrak{T}_1$ durch C und $\mathfrak{T}_2$ durch und vermindern wir hinreichend da Variabilitätsintervall des Parameter der Schar, so können wir zweierlei e reichen: 1. der Bogen $C\mathfrak{E}D$ wird sic in einem von zwei Transversalen $\mathfrak{T}$ und $\mathfrak{T}_2$ und zwei Grenzextremale begrenzten krummlinigen Viereck befinden, der vom Extremalenfelde erfül ist und in welchem jeder Bogen, der seine Endpunkte auf $\mathfrak{T}_1$ und $\mathfrak{T}_2$ ha einen kleineren Wert für J liefert, als $C\mathfrak{E}D$; 2. die Bogen $A\mathfrak{E}B$ un $D\mathfrak{E}B$ werden keine gemeinsamen Punkte mit den Grenzextremalen habe Mit F bezeichnen wir das Innere des Vierecks.

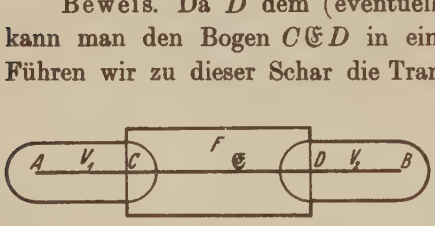


Fig. 1.

Sei ϱ eine positive Zahl, die kleiner ist als die kleinste Entfernu der Bogen $A\mathfrak{E}C$ und $D\mathfrak{E}B$ von den beiden Grenzextremalen und als d Hälfte der kleinsten Entfernung zwischen $A\mathfrak{E}C$ und $D\mathfrak{E}B$. Führen w die Umgebungen $V_1 = U_\varrho(A\mathfrak{E}C)$ und $V_2 = U_\varrho(D\mathfrak{E}B)$ ein. Die Vereinigung menge von V_1, F und V_2 bildet eine Umgebung U von $A\mathfrak{E}C$, die, wie w zeigen werden, den Forderungen des Hilfssatzes entspricht.

In der Tat, sei $\mathfrak{L}$ eine Kurve, die C und D verbindet und in d Umgebung U liegt. Sei (6) ihre Parameterdarstellung. Die Kurve $\mathfrak{L}$ h mit der Begrenzung von F die Punkte C und D gemeinsam. Hat sie noc andere Punkte, so müssen das innere Punkte von V_1 oder V_2 sein, also z

$\mathfrak{T}_1$ oder $\mathfrak{T}_2$ gehören (die Grenzextremalen haben nach der Wahl von ϱ eine gemeinsamen Punkte mit V_1 und V_2). Sei τ_1 der größte Parameterwert, der einem mit $\mathfrak{T}_1$ gemeinsamen Punkte von $\mathfrak{L}$ entspricht, τ_2 der kleinste, der größer als τ_1 ist und der einem mit $\mathfrak{T}_2$ gemeinsamen Punkte entspricht. Dann ist

$$a \leqq \tau_1 < \tau_2 \leqq b .$$

Der Anfangspunkt des Bogens von $\mathfrak{L}$, dem die Werte $\tau_1 \leqq \tau \leqq \tau_2$ entsprechen, liegt auf $\mathfrak{T}_1$, also auch in V_1, sein Endpunkt auf $\mathfrak{T}_2$ und in V_2. Außerdem hat er keine gemeinsamen Punkte mit der Begrenzung von F. Da die Umgebungen V_1 und V_2 nach der Wahl von ϱ getrennt sind, so muß der erwähnte Bogen ganz in F liegen.

Ihm entspricht ein Wert des Integrals J, der nicht kleiner ist als $J(C\mathfrak{E}D)$. Wegen (3) ist auch längs der ganzen Kurve $\mathfrak{L}$

$$J(C\mathfrak{L}D) \geqq J(C\mathfrak{E}D).$$

Das Gleichheitszeichen tritt nur dann auf, wenn $\tau_1 = a$ und $\tau_2 = b$ und wenn außerdem die Kurve $C\mathfrak{L}D$ mit einer Extremalen, also mit $C\mathfrak{E}D$ zusammenfällt.

Fällt ein oder beide Endpunkte C und D mit A oder B zusammen, so bleibt der Beweis nach einigen Vereinfachungen gültig.

Damit ist unser Hilfssatz bewiesen worden.

4. Hinreichende Bedingung.

Satz. *Enthält die Extremale $\mathfrak{E}$, die die Punkte A und B verbindet, den zweiten konjugierten Punkt von A nicht, so besitzt sie die Minimaxeigenschaft.*

Die ausgesprochene Bedingung kann man so fassen

$$B \prec A^{(2)},$$

indem man (7) als erfüllt ansieht, wenn $A^{(3)}$ überhaupt nicht existiert.

Beweis. Nach (7) geht $B^{(-1)}$ auf der Extremale $\mathfrak{E}$ dem Punkte $A^{(1)}$ voran (Fig. 2). Wählen wir zwei Punkte C und D zwischen $B^{(-1)}$ und $A^{(1)}$.

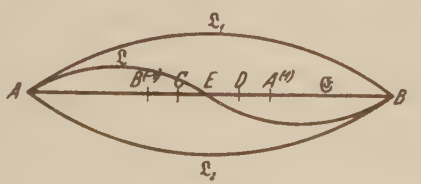


Fig. 2.

Nach dem Hilfssatz des § 3 können wir für die Extremale $A\mathfrak{E}B$ eine Umgebung bilden, in der der Bogen $A\mathfrak{E}D$ ein Minimum liefert, und eine zweite Umgebung, in der dasselbe den Bogen $C\mathfrak{E}B$ betrifft. Die Durchschnittsmenge U der beiden Umgebungen ist auch eine Umgebung von $A\mathfrak{E}B$, in der die beiden Bogen $A\mathfrak{E}D$ und $C\mathfrak{E}B$ und alle ihre Teilbogen ein Minimum des Integrals J liefern.

Seien $\mathfrak{L}_1$ und $\mathfrak{L}_2$ zwei Kurven, die die Punkte A und B verbinden die in U auf beiden Seiten von $A\mathfrak{E}B$ liegen, und längs deren J kleiner als $J(A\mathfrak{E}B)$ ist. Betrachten wir eine *beliebige* stetige Überführung der Kurve $\mathfrak{L}_1$ in $\mathfrak{L}_2$ innerhalb U. Die Übergangskurve $\mathfrak{L}$ setzen wir mit $\overleftarrow{\mathfrak{L}_2}$ zu einer geschlossenen Kurve $\mathfrak{D}$ zusammen. Für einen Parameterwert ist $\mathfrak{D}$ aus $\mathfrak{L}_1$ und $\overleftarrow{\mathfrak{L}_2}$ zusammengesetzt, also nach der Definition 1 in bezug auf $(P - C\mathfrak{E}D)$ irreduzibel. Für einen anderen ist $\mathfrak{D}$ aus $\mathfrak{L}_2$ und $\overleftarrow{\mathfrak{L}_2}$ zusammengesetzt, also in bezug auf $(P - C\mathfrak{E}D)$ reduzibel. Infolgedessen kann $\mathfrak{D}$ nicht immer ganz in $(P - C\mathfrak{E}D)$ liegen und muß für gewisse Parameterwerte mit $C\mathfrak{E}D$ gemeinsame Punkte besitzen. Diese gehören selbstverständlich dem Bestandteile $\mathfrak{L}$ von $\mathfrak{D}$ an.

Sei E ein solcher Punkt von $C\mathfrak{E}D$, der einer gewissen Kurve $\mathfrak{L}$ angehört. Der Eigenschaft der Umgebung U halber haben wir

$$J(A\mathfrak{L}E) \geqq J(A\mathfrak{E}E),$$
$$J(E\mathfrak{L}B) \geqq J(E\mathfrak{E}B),$$

und nach Addition

$$J(A\mathfrak{L}B) \geqq J(A\mathfrak{E}B).$$

Also, bei beliebiger Überführung der Kurve $\mathfrak{L}_1$ in $\mathfrak{L}_2$ muß man auf die Kurven $\mathfrak{L}$ stoßen, längs deren das Integral J nicht kleiner als längs $A\mathfrak{E}B$ ist. Mit anderen Worten gesagt, besitzt die Extremale $A\mathfrak{E}B$ die Minimaxeigenschaft.

5. Notwendige Bedingung.

Satz. *Besitzt die Extremale $\mathfrak{E}$, die die Punkte A und B verbinden, die Minimaxeigenschaft, so geht der zweite konjugierte Punkt von A den Punkte B nicht voran.*

Das heißt, daß aus der Minimaxeigenschaft folgt, daß

$$B \leqq A^{(2)}. \tag{8}$$

Beweis. Setzen wir im Gegenteil voraus, daß

$$A^{(2)} < B. \tag{9}$$

Sei U eine *beliebige* Umgebung von $A\mathfrak{E}B$.

Man kann von A und von B zwei unendliche Bogen führen, die in A und in B von den Tangenten des Bogens $A\mathfrak{E}B$ verschiedene Tangenten besitzen und die mit $A\mathfrak{E}B$ zusammen einen unendlichen Bogen ohne mehrfache Punkte bilden, der die Ebene P in zwei Teile zerschneidet, die wir positive und negative Seite von $A\mathfrak{E}B$ nennen. Liegen zwei Kurven die A und B verbinden, eine auf der positiven, die andere auf der negativen Seite von $A\mathfrak{E}B$, so sieht man leicht, daß sie auf beiden Seiten (im Sinne der Definition 1 des § 2) von $A\mathfrak{E}B$ liegen.

Nach (9) geht $A^{(1)}$ dem Punkte $B^{(-1)}$ voran (Fig. 3). Wählen wir auf $\mathfrak{E}$ zwischen $A^{(1)}$ und $B^{(-1)}$ zwei Punkte C und D. Die Bogen $A\mathfrak{E}C$ und $D\mathfrak{E}B$ liefern kein Minimum des Integrals J. In beliebiger Umgebung von $A\mathfrak{E}E$ können wir Kurven bilden, die A und C, beziehungsweise D und B verbinden und längs deren das Integral J kleinere Werte annimmt als längs $A\mathfrak{E}C$, beziehungsweise $D\mathfrak{E}B$. Wir können solche Kurven an beiden Seiten von $A\mathfrak{E}B$ bilden. Es genügt dazu, sich der Beweise von Herrn Hadamard[3]) oder vom Verfasser[4]) für die Notwendigkeit der Jacobischen Bedingung zu bedienen. Aus denselben Beweisen folgt, daß man A und C, beziehungsweise D und B mit ganzen einparametrigen Scharen solcher Kurven verbinden kann, die stetig von einem Parameter abhängen und für einen speziellen Wert des Parameters mit $A\mathfrak{E}C$, beziehungsweise mit $D\mathfrak{E}B$ zusammenfallen.

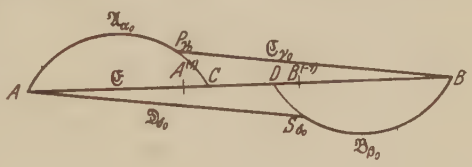


Fig. 3.

Bilden wir eine einparametrige Schar der Kurven

$$\mathfrak{A}_\alpha\colon\quad x = x_1(\tau, \alpha), \qquad y = y_1(\tau, \alpha), \qquad 0 \leqq \tau \leqq \tau_1(\alpha),$$

die folgende Eigenschaften hat:

1. Für jedes α des Intervalles $0 \leqq \alpha \leqq \alpha_0$ verbindet die Kurve $\mathfrak{A}_\alpha$ die Punkte A und C.

2. Für $\alpha = 0$ fällt die Kurve $\mathfrak{A}_\alpha = \mathfrak{A}_0$ mit $A\mathfrak{E}C$ zusammen.

3. Für $0 < \alpha \leqq \alpha_0$ liegt $\mathfrak{A}_\alpha$ in U auf der *positiven* Seite von $A\mathfrak{E}B$.

4. Die Funktionen $x_1(\tau, \alpha)$ und $y_1(\tau, \alpha)$ sind nebst ihren ersten Ableitungen nach τ stetig für $0 \leqq \alpha \leqq \alpha_0$, $0 \leqq \tau \leqq \tau_1(\alpha)$.

5. Für $0 < \alpha \leqq \alpha_0$ ist

$$J(A\mathfrak{A}_\alpha C) < J(A\mathfrak{E}C).$$

Selbstverständlich hätte hier für $\alpha = 0$ ein Gleichheitszeichen Platz. Analog bilden wir eine zweite einparametrige Kurvenschar

$$\mathfrak{B}_\beta\colon\quad x = x_2(\tau, \beta), \qquad y = y_2(\tau, \beta), \qquad 0 \leqq \tau \leqq \tau_2(\beta)$$

mit folgenden Eigenschaften:

1. Für jedes β des Intervalles $0 \leqq \beta \leqq \beta_0$ verbindet die Kurve $\mathfrak{B}_\beta$ die Punkte D und B.

2. Für $\beta = 0$ fällt die Kurve $\mathfrak{B}_\beta = \mathfrak{B}_0$ mit $D\mathfrak{E}B$ zusammen.

3. Für $0 < \beta \leqq \beta_0$ liegt $\mathfrak{B}_\beta$ in U auf der *negativen* Seite von $A\mathfrak{E}B$.

---

[3]) Leçons sur le Calcul des Variations **1** (1910), S. 402.

[4]) Math. Annalen **101** (1929), S. 670.

4. Die Funktionen $x_2(\tau, \beta)$ und $y_2(\tau, \beta)$ sind nebst ihren ersten Ableitungen nach τ stetig für $0 \leqq \beta \leqq \beta_0$, $0 \leqq \tau \leqq \tau_2(\beta)$.

5. Für $0 < \beta \leqq \beta_0$ ist

$$J(D\mathfrak{B}_\beta B) < J(D\mathfrak{E}B).$$

Auch hier hätte für $\beta = 0$ das Gleichheitszeichen Platz.

Die Konstruktion der beiden Kurvenscharen ist nach dem oben Gesagten möglich.

Jetzt bilden wir eine dritte Kurvenschar

$$\mathfrak{C}_\gamma\colon\ x = x_3(\tau, \gamma), \qquad y = y_3(\tau, \gamma), \qquad 0 \leqq \tau \leqq \tau_3(\gamma)$$

mit folgenden Eigenschaften:

1. Für jedes γ eines Intervalles $0 \leqq \gamma \leqq \gamma_0$ verbindet die Kurve $\mathfrak{C}_\gamma$ einen Punkt P_γ der Kurve $\mathfrak{A}_{\alpha_0}$, dem der Parameterwert $\tau = \tau_1(\alpha_0) - \gamma$ entspricht, mit dem Punkte B.

2. Für $\gamma = 0$ fällt die Kurve $\mathfrak{C}_\gamma = \mathfrak{C}_0$ mit $C\mathfrak{E}B$ zusammen.

3. Für $0 < \gamma \leqq \gamma_0$ liegt $\mathfrak{C}_\gamma$ in U an *positiver* Seite von $A\mathfrak{E}B$.

4. Die Funktionen $x_3(\tau, \gamma)$ und $y_3(\tau, \gamma)$ sind nebst ihren ersten Ableitungen nach τ stetig für $0 \leqq \gamma \leqq \gamma_0$, $0 \leqq \tau \leqq \tau_3(\gamma$.

Endlich bilden wir eine vierte Kurvenschar

$$\mathfrak{D}_\delta\colon\ x = x_4(\tau, \delta), \qquad y = y_4(\tau, \delta), \qquad 0 \leqq \tau \leqq \tau_4(\delta)$$

mit folgenden Eigenschaften:

1. Für jedes δ eines Intervalles $0 \leqq \delta \leqq \delta_0$ verbindet die Kurve $\mathfrak{D}_\delta$ den Punkt A mit einem Punkte S_δ der Kurve $\mathfrak{B}_{\beta_0}$, dem der Parameterwert $\tau = \delta$ entspricht.

2. Für $\delta = 0$ fällt die Kurve $\mathfrak{D}_\delta = \mathfrak{D}_0$ mit $A\mathfrak{E}D$ zusammen.

3. Für $0 < \delta \leqq \delta_0$ liegt $\mathfrak{D}_\delta$ in U an *negativer* Seite von $A\mathfrak{E}B$.

4. Die Funktionen $x_4(\tau, \delta)$ und $y_4(\tau, \delta)$ sind nebst ihren ersten Ableitungen nach τ stetig für $0 \leqq \delta \leqq \delta_0$, $0 \leqq \tau \leqq \tau_4(\delta)$.

Das Integral J ist längs der Kurve $A\mathfrak{A}_{\alpha_0}C\mathfrak{E}B$ kleiner als $J(A\mathfrak{E}B)$. Wir wählen γ_0 hinreichend klein, so daß J auch längs der Kurve $A\mathfrak{A}_{\alpha_0}P_\gamma\mathfrak{C}_\gamma B$ für $0 \leqq \gamma \leqq \gamma_0$ kleiner als $J(A\mathfrak{E}B)$ sei. In ähnlicher Weise wählen wir δ_0 hinreichend klein, so daß J längs der Kurven $A\mathfrak{D}_\delta S_\delta \mathfrak{B}_{\beta_0}B$ für $0 \leqq \delta \leqq \delta_0$ kleiner als $J(A\mathfrak{E}B)$ sei.

Jetzt gehen wir zum Beweise über. Wir bezeichnen die Kurven $A\mathfrak{A}_{\alpha_0}P_{\gamma_0}\mathfrak{C}_{\gamma_0}B$ und $A\mathfrak{D}_{\delta_0}S_{\delta_0}\mathfrak{B}_{\beta_0}B$ mit $\mathfrak{L}_p$ und $\mathfrak{L}_n$. Dann gelten folgende Behauptungen:

1. Die Kurve $\mathfrak{L}_p$ liegt in U auf der positiven, $\mathfrak{L}_n$ auf der negativen, also auf verschiedenen Seiten von $A\mathfrak{E}B$.

2. Das Integral J nimmt längs $\mathfrak{L}_p$ und $\mathfrak{L}_n$ einen kleineren Wert an als $J(A\mathfrak{E}B)$.

3. Man kann $\mathfrak{L}_P$ in $\mathfrak{L}_n$ stetig in U so überführen, daß längs aller Übergangskurven J kleiner als $J(A\mathfrak{E}B)$ ist.

Die Punkte 1. und 2. folgen aus der Konstruktion der Scharen und der Wahl von γ_0 und δ_0. Wir sollen nur kurz zeigen, wie man den in 3. erwähnten Übergang vollenden soll. Zu diesem Zweck:

a) Betrachten wir die Kurve $A\mathfrak{A}_{\alpha_0}P_{\gamma}\mathfrak{C}_{\gamma}B$ und verändern γ von γ_0 bis 0. Für $\gamma=\gamma_0$ fällt diese Kurve mit $\mathfrak{L}_p$ zusammen, für $\gamma=0$ mit $A\mathfrak{A}_{\alpha_0}C\mathfrak{E}B$.

b) Jetzt betrachten wir die Kurve $A\mathfrak{A}_{\alpha_0}C\mathfrak{E}D\mathfrak{B}_{\beta}B$ und verändern β von 0 bis β_0. Für $\beta=0$ fällt diese Kurve mit der Endkurve von a) zusammen, für $\beta=\beta_0$ mit $A\mathfrak{A}_{\alpha_0}C\mathfrak{E}D\mathfrak{B}_{\beta_0}B$.

c) Jetzt betrachten wir die Kurve $A\mathfrak{A}_{\alpha}C\mathfrak{E}D\mathfrak{B}_{\beta_0}B$ und verändern α von α_0 bis 0. Für $\alpha=\alpha_0$ fällt diese Kurve mit der Endkurve von b) zusammen, für $\alpha=0$ mit $A\mathfrak{E}D\mathfrak{B}_{\beta_0}B$.

d) Endlich betrachten wir die Kurve $A\mathfrak{T}_{\delta}S_{\delta}\mathfrak{B}_{\beta_0}B$ und verändern δ von 0 bis δ_0. Für $\delta=0$ fällt diese Kurve mit der Endkurve von c) zusammen, für $\delta=\delta_0$ mit $\mathfrak{L}_n$.

Man kann leicht bestätigen, daß dieser Übergang alle Forderungen erfüllt.

Da in beliebiger Umgebung U von $A\mathfrak{E}B$ die Kurven $\mathfrak{L}_p$ und $\mathfrak{L}_n$ existieren, für die die oben aufgezählten Behauptungen gelten, hat die Extremale $A\mathfrak{E}B$ keine Minimaxeigenschaft. Infolgedessen muß nicht (9), sondern (8) gelten, was zu beweisen war.

6. Bemerkung.

Die Sätze von Paragraphen 4 und 5 lassen den Fall unentschieden, in welchem der Punkt $A^{(2)}$ mit B zusammenfällt.

Zweites Kapitel.

Geschlossene Extremalen.

7. Voraussetzungen und Definitionen.

Jetzt soll $\mathfrak{E}$ eine geschlossene, nichtgebrochene Extremale ohne mehrfache Punkte bedeuten. Ihre Länge sei d. Wählt man, wie vorher, die Bogenlänge t als Parameter, so kann (2) wieder $\mathfrak{E}$ darstellen. Doch kann man das Veränderungsintervall von t auf die ganze t-Achse ausdehnen und die Funktionen $\varphi(t)$ und $\psi(t)$ periodisch machen:

$$\varphi(t+d)=\varphi(t), \qquad \psi(t+d)=\psi(t). \tag{10}$$

Die Voraussetzungen (3), (4), (5) sollen auch jetzt erfüllt sein. Auch R soll so vermindert werden, daß (3) im ganzen R erfüllt sei.

Definition 1. *Wir sagen, daß eine Kurve $\mathfrak{L}$ der Umgebung U_ϱ der Kurve $\mathfrak{E}$ ordentlich (italienisch: ordinatamente) angehört, wenn man zwischen den Punkten von $\mathfrak{L}$ und $\mathfrak{E}$ eine solche Zuordnung einrichten kann, daß*

1. *jedem Punkte einer Kurve entweder ein Punkt oder ein Teilbogen der anderen entspricht,*

2. *entsprechen einander zwei Tripel nicht zusammenfallender Punkte der beiden Kurven, so entspricht der positiven Anordnung eines Tripels die positive Anordnung des zweiten,*

3. *die Entfernung der zugeordneten Punkte kleiner als ϱ ist*[5]).

Gehört $\mathfrak{L}$ der Umgebung U_ϱ ordentlich, so liegt $\mathfrak{L}$ in U_ϱ.

Die Zuordnung, von der in der Definition 1 die Rede ist, werden wir eine ϱ-Zuordnung nennen. Selbstverständlich können zwischen $\mathfrak{L}$ und $\mathfrak{E}$ viele ϱ-Zuordnungen stattfinden. Eine ϱ-Zuordnung bestimmt eine Zuordnung der Sinne der beiden Kurven. Man könnte sich überzeugen, daß für hinreichend kleine ϱ alle ϱ-Zuordnungen dieselbe Zuordnung der Sinne der Kurven bestimmen.

Die Extremale $\mathfrak{E}$ zerschneidet die Ebene P in zwei Teile, die wir nach Belieben *positive und negative Seite von* $\mathfrak{E}$ nennen.

Definition 2. *Wir sagen, daß die geschlossene Extremale $\mathfrak{E}$ die Minimaxeigenschaft besitzt, wenn für eine hinreichend kleine positive Zahl ϱ die folgende Bedingung erfüllt ist:*

Wenn es zwei geschlossene Kurven gibt, die der Umgebung $U_\varrho(\mathfrak{E})$ ordentlich angehören, auf beiden Seiten von $\mathfrak{E}$ liegen und längs deren das Integral J einen kleineren Wert annimmt als längs $\mathfrak{E}$, so lassen sich diese Kurven nicht ineinander stetig in der Klasse der Kurven überführen, die der Umgebung $U_\varrho(\mathfrak{E})$ ordentlich angehören und längs deren das Integral J kleiner bleibt als längs $\mathfrak{E}$.

Auch jetzt umfaßt unsere Definition die Fälle des Minimums und des einseitigen Minimums.

8. Hilfssatz.

Sei t ein beliebiger Parameterwert auf $\mathfrak{E}$. *Mit t^* bezeichnen wir den kleineren von den zwei Werten $t^{(1)}$ und $t + d$.* Entspricht t einem Punkte P, so ist t^* einem Punkte P^* zugeordnet, den wir *den reduzierten ersten*

[5]) Tonelli, Calcolo delle Variazioni 1 (1921), S. 72. Was den Punkt 2. betrifft, so haben wir nur die Definition von Herrn Tonelli für die Betrachtung der geschlossenen Kurven geeignet gemacht.

konjugierten Punkt nennen. Dieser fällt entweder mit P selbst oder mit $P^{(1)}$ zusammen, je nachdem der Bogen $P\mathfrak{E}P^{(1)}$ die ganze Extremale $\mathfrak{E}$ enthält oder nicht[6]). Da $t^{(1)}$ und $t+d$ beide monoton wachsende stetige Funktionen von t sind, so gilt dasselbe für die Funktion t^* von t.

Sei ε eine beliebig kleine positive Zahl, die vom Minimum der stetigen Funktion t^*-t von t im Intervalle $0 \leqq t \leqq d$ kleiner ist. Da diese Funktion periodisch ist, so haben wir für jedes t

$$\varepsilon < t^* - t. \tag{11}$$

Ich behaupte, daß es für eine solche Zahl ε eine solche positive Zahl ϱ gibt, daß,

wenn $C\mathfrak{E}D$ ein beliebiger Bogen ist, für welchen D mehr als um ε dem Punkte C^* vorangeht,

wenn C^- dem Punkte C um $\frac{\varepsilon}{3}$ vorangeht und D^+ dem Punkte D um $\frac{\varepsilon}{3}$ nachfolgt,

so liefert der Bogen $C\mathfrak{E}D$ einen kleineren Wert für das Integral J, als jeder andere Bogen, der C und D verbindet und in der Umgebung $U_\varrho(C^-\mathfrak{E}D^+)$ liegt.

Es ist zu bemerken, daß aus (11) die Existenz der Bogen $C\mathfrak{E}D$ folgt und daß der Bogen $C^-\mathfrak{E}D^+$ nicht die ganze Extremale $\mathfrak{E}$ enthält.

Beweis. Da die Funktion t^* von t stetig ist, kann man die Extremale $\mathfrak{E}$ so in Teilbogen $A_0\mathfrak{E}A_1$, $A_1\mathfrak{E}A_2, \ldots, A_{n-1}\mathfrak{E}A_0$ einteilen, daß die Länge jedes Teilbogens, wie auch jedes der entsprechenden Bogen $A_0^*\mathfrak{E}A_1^*$, $A_1^*\mathfrak{E}A_2^*, \ldots, A_{n-1}^*\mathfrak{E}A_0^*$ kleiner als $\frac{\varepsilon}{8}$ ist. Mit B_k $(k=0,1,\ldots,n-1)$ bezeichnen wir Punkte, die den Punkten A_k^* um $\frac{3}{4}\varepsilon$ vorangehen. Nach (11) folgt B_k dem Punkte A_k nach. Mit A_k^- bezeichnen wir den Punkt, der dem Punkte A_k um $\frac{\varepsilon}{3}$ vorangeht, mit B_k^+ den, der dem Punkte B_k um $\frac{\varepsilon}{3}$ nachfolgt. Der Bogen $\bar{A}_k\mathfrak{E}\bar{B}_k$ enthält nicht die ganze Extremale $\mathfrak{E}$.

Nach dem Hilfssatze des § 3 gibt es für jeden Bogen $A_k\mathfrak{E}B_k$ eine solche positive Zahl ϱ_k, daß in der Umgebung $U_{\varrho_k}(A_k^-\mathfrak{E}B_k^+)$ der Bogen $A_k\mathfrak{E}B_k$ ein Minimum liefert. Dasselbe trifft für jeden Teilbogen von $A_k\mathfrak{E}B_k$ zu. Sei ϱ die kleinste von den Zahlen ϱ_k.

Betrachten wir einen beliebigen Bogen $C\mathfrak{E}D$, wo D dem Punkte C^* mehr als um ε vorangeht. Mit $\bar{C}$ bezeichnen wir den Punkt, der dem Punkte C um $\frac{\varepsilon}{3}$ vorangeht, mit $\bar{D}$ den, der dem Punkte D um $\frac{\varepsilon}{3}$ nach-

[6]) Die Bogen der Extremale $\mathfrak{E}$ werden immer im positiven Sinne durchlaufen. So bedeutet $P\mathfrak{E}P^{(1)}$ denjenigen von den zwei Bogen der Extremale $\mathfrak{E}$, auf dem der Punkt P dem $P^{(1)}$ vorangeht.

folgt. Der Punkt C liegt auf einem gewissen Bogen $A_k \mathfrak{E} A_{k+1}$.[7]) Der Punkt C^* liegt auf dem Bogen $A_k^* \mathfrak{E} A_{k+1}^*$, geht also dem Punkte A_{k+1}^* voran. Der Punkt D geht dem Punkte C^*, also auch dem Punkte A_{k+1}^* mehr als um ε voran. Der Punkt A_k^* geht dem Punkte A_{k+1}^* weniger als um $\frac{\varepsilon}{8}$, und B_k dem Punkte A_k^* genau um $\frac{3}{4}\varepsilon$ voran. Also geht B_k dem Punkte A_{k+1}^* weniger als um $\frac{\varepsilon}{8} + \frac{3}{4}\varepsilon < \varepsilon$ voran. Daraus folgt, daß D dem Punkte B_k vorangeht, und daß der Bogen $C \mathfrak{E} D$ in $A_k \mathfrak{E} B_k$ enthalten ist. Auch der Bogen $C^- \mathfrak{E} D^+$ ist in $A_k^- \mathfrak{E} B_k^+$ enthalten.

Die Umgebung $U_\varrho(C^- \mathfrak{E} D^+)$ ist in $U_\varrho(A_k^- \mathfrak{E} B_k^+)$ und daher auch in $U_{\varrho_k}(A_k^- \mathfrak{E} B_k^+)$ enthalten. Da in der letzten Umgebung der Bogen $A_k \mathfrak{E} B_k$, wie auch jeder Teilbogen $C \mathfrak{E} D$ Minimum liefert, so liefert $C \mathfrak{E} D$ Minimum auch in $U_\varrho(C^- \mathfrak{E} D^+)$.

Damit ist der Hilfssatz bewiesen.

9. Hinreichende Bedingung.

Satz. *Enthält die geschlossene Extremale $\mathfrak{E}$ den zweiten konjugierten Punkt keines ihrer Punkte, so besitzt sie die Minimaxeigenschaft.*

Mit anderen Worten, setzen wir voraus, daß für jedes t

$$t + d < t^{(2)}. \tag{12}$$

Wir zerspalten die Beweisführung in einzelne Punkte.

a) Die Zahl ε. Wir führen jetzt die Funktion $t^{**} = (t^*)^*$ ein. Sie ist eine stetige, monoton-wachsende Funktion von t. Ist $t^* = t + d$ oder $t^{**} = t^* + d$, so ist sicher $t^{**} > t + d$. Ist dagegen $t^* = t^{(1)}$ und $t^{**} = (t^*)^{(1)}$, so ist $t^{**} = t^{(2)}$ und nach (12) $t^{**} > t + d$. Also in jedem Falle haben wir

$$t^{**} > t + d.$$

Sei 2δ eine positive Zahl, die kleiner ist als das Minimum der stetigen und positiven Funktion $t^{**} - t - d$ im Intervalle $0 \leqq t \leqq d$. Da diese Funktion periodisch ist, so haben wir für jedes t

$$t^{**} - t > d + 2\delta. \tag{13}$$

Sei ε eine positive Zahl, für die

1. $2\varepsilon < \delta$,

2. 4ε kleiner als das Minimum der stetigen Funktion $t^* - t$ im Intervalle $0 \leqq t \leqq d$ ist,

3. ist der Zuwachs von t kleiner als 2ε, so ist der Zuwachs von t^{**} kleiner als δ.

---

[7]) Für $k = n - 1$ anstatt $k + 1 = n$ ist der Wert 0 zu setzen.

Aus 1. folgt, daß für jedes t

$$t^* - t > 4\varepsilon \tag{14}$$

und daraus, daß

$$d > 4\varepsilon. \tag{15}$$

Aus (14) folgt, daß die Zahl ε die Bedingung (11) erfüllt, und daß man eine Zahl ϱ finden kann, für die der Hilfssatz des § 8 gilt.

b) Die Zahl ϱ. In der Tat bilden wir eine solche positive Zahl ϱ, daß

1. für ϱ die Minimumseigenschaft, von der in dem Hilfssatze die Rede ist, gilt,

2. 3ϱ kleiner als Minimum der Chorde eines Bogens von $\mathfrak{E}$ ist, dessen Länge l die Ungleichheit $\frac{\varepsilon}{3} \leqq l \leqq d - \frac{\varepsilon}{3}$ erfüllt.

Wir werden beweisen, daß für die Zahl ϱ die Minimaxeigenschaft der Definition 2 des § 7 gemäß gilt.

c) Die Klassen $\mathfrak{A}$ und $\mathfrak{B}$. Im folgenden werden wir immer mit $\mathfrak{L}$ (auch mit Indizes) nur solche geschlossene Kurven bezeichnen, die zu $U_\varrho(\mathfrak{E})$ ordentlich gehören, und längs deren das Integral J einen kleineren Wert annimmt als $J(\mathfrak{E})$.

Sei $\mathfrak{L}$ eine solche Vergleichskurve. Die Menge der gemeinsamen Punkte von $\mathfrak{E}$ und $\mathfrak{L}$ ist auf $\mathfrak{E}$ abgeschlossen. Die komplementäre Menge auf $\mathfrak{E}$ ist offen und besteht aus höchstens abzählbar vielen offenen Intervallen σ_k. Seien die Endpunkte des Intervalles σ_k A_k und B_k. Wir ordnen diesen Punkten die Parameterwerte t_k und $\bar{t}_k$ so, daß $t_k < \bar{t}_k$ und $\bar{t}_k - t_k < d$.

Wir teilen alle Kurven $\mathfrak{L}$ in zwei Klassen $\mathfrak{A}$ und $\mathfrak{B}$ ein:

1. *Die Kurve $\mathfrak{L}$ gehört zur Klasse $\mathfrak{A}$, wenn entweder mindestens für ein Intervall σ_k die Ungleichheit $\bar{t}_k > t_k^* - 2\varepsilon$ gilt oder die Kurve $\mathfrak{L}$ keinen gemeinsamen Punkt mit $\mathfrak{E}$ hat.*

2. *Die Kurve $\mathfrak{L}$ gehört zur Klasse $\mathfrak{B}$, wenn für jedes Intervall σ_k die Ungleichheit $\bar{t}_k \leqq t_k^* - 2\varepsilon$ gilt.*

Man sieht leicht, daß im Falle der Klasse $\mathfrak{A}$ *die Ungleichheit $\bar{t}_k > t_k^* - 2\varepsilon$ nur für ein Intervall σ_k erfüllt sein kann.* In der Tat, setzen wir voraus, daß es zwei solche Intervalle σ_k und σ_l gibt. Man kann t_k und t_l so wählen, daß

$$t_k < \bar{t}_k \leqq t_l < \bar{t}_l \leqq t_k + d \tag{16}$$

und

$$\bar{t}_k > t_k^* - 2\varepsilon, \tag{17a}$$

$$\bar{t}_l > t_l^* - 2\varepsilon. \tag{17b}$$

Aus (16) und (17a) folgt

$$t_l > t_k^* - 2\varepsilon.$$

Daraus nach dem Punkte 3. der Definition von ε

$$t_l^* > t_k^{**} - \delta.$$

Jetzt nach (16) und (17b)

$$t_k + d > t_k^{**} - \delta - 2\varepsilon$$

oder nach dem Punkte 1. der Definition von ε

$$t_k + d > t_k^{**} - 2\delta,$$

was der Ungleichheit (13) widerspricht. Damit ist das Geforderte bewiesen worden.

d) Die Klassen $\mathfrak{A}_I$ und $\mathfrak{A}_{II}$. Wir führen jetzt eine Einteilung der Kurven $\mathfrak{L}$ der Klasse $\mathfrak{A}$ in zwei Unterklassen $\mathfrak{A}_I$ und $\mathfrak{A}_{II}$ ein.

Sei $\mathfrak{L}$ eine Kurve der Klasse $\mathfrak{A}$, die mit $\mathfrak{E}$ gemeinsame Punkte hat. Sei $A_k \mathfrak{E} B_k$ der einzige Bogen σ_k, für welchen $\bar{t}_k > t_k^* - 2\varepsilon$. Nach (14) haben wir

$$\bar{t}_k > t_k + 2\varepsilon. \tag{18}$$

Betrachten wir die Punkte der Extremale $\mathfrak{E}$ mit entsprechenden Werten von t (Fig. 4): A_k mit $t = t_k$, C_k mit $t = t_k + \frac{\varepsilon}{3}$, E_k mit $t = t_k + \frac{2\varepsilon}{3}$, F_k mit $t = \bar{t}_k - \frac{2\varepsilon}{3}$, D_k mit $t = \bar{t}_k - \frac{\varepsilon}{3}$ und B_k mit $t = \bar{t}_k$. Nach (18) folgt F_k dem Punkte E_k nach. Wir führen die Bezeichnung $V_\varrho = U_\varrho(C_k \mathfrak{E} D_k)$, $W_\varrho = U_\varrho(E_k \mathfrak{E} F_k)$ ein. Selbstverständlich ist dann $W_\varrho \prec V_\varrho \prec U_\varrho(\mathfrak{E})$ (in der Figur sind diese Umgebungen mit stetiger Linie gezeichnet).

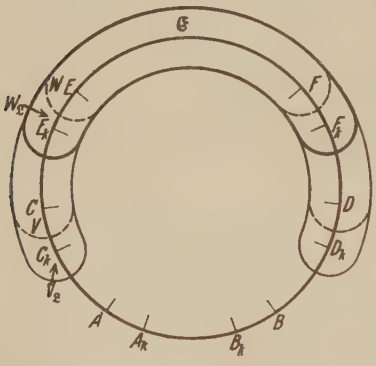


Fig. 4.

Bei beliebiger ϱ-Zuordnung wird dem Bogen $C_k \mathfrak{E} D_k$ ein Bogen $M_k \mathfrak{L} N_k$ entsprechen, der ganz in V_ϱ liegt. Da die Bogen $A_k \mathfrak{E} C_k$ und $D_k \mathfrak{E} B_k$ von der Länge $\frac{\varepsilon}{3}$ sind, so ist die Entfernung jedes Punktes des Bogens $C_k \mathfrak{E} D_k$ von $B_k \mathfrak{E} A_k$ größer als 3ϱ (vgl. den Punkt 2. der Definitionvon ϱ). Infolgedessen ist die Entfernung jedes Punktes der Umgebung V_ϱ von $B_k \mathfrak{E} A_k$ größer als 2ϱ: der Bogen $M_k \mathfrak{L} N_k$ hat mit $B_k \mathfrak{E} A_k$ keinen gemeinsamen Punkt. Da es im Innern von $A_k \mathfrak{E} B_k$ überhaupt keine Punkte von $\mathfrak{L}$ gibt, so kann $M_k \mathfrak{L} N_k$ keine gemeinsamen Punkte mit $\mathfrak{E}$ haben und muß ganz an einer Seite von $\mathfrak{E}$ liegen.

Betrachten wir jetzt die Menge M_ϱ der Punkte von $\mathfrak{L}$, die in W_ϱ liegen. Diese kann nicht leer sein, weil sie den Bogen enthält, der bei beliebiger ϱ-Zuordnung dem Bogen $E_k \mathfrak{E} F_k$ entspricht. Da die Länge der

Bogen $C_k \mathfrak{E} E_k$ und $F_k \mathfrak{E} D_k$ gleich $\frac{\varepsilon}{3}$ ist, so ist die Entfernung jedes Punktes des Bogens $E_k \mathfrak{E} F_k$ von $D_k \mathfrak{E} C_k$ größer als 3ϱ. Infolgedessen ist die Entfernung jedes Punktes der Umgebung W_ϱ von $D_k \mathfrak{E} C_k$ größer als 2ϱ. Also kann ein Punkt von M_ϱ bei keiner ϱ-Zuordnung den Punkten von $D_k \mathfrak{E} C_k$ entsprechen. Er muß einem Punkte oder einem Intervalle von $C_k \mathfrak{E} D_k$ zugeordnet sein und muß damit dem Bogen $M_k \mathfrak{L} N_k$ angehören, unabhängig davon, was für eine ϱ-Zuordnung aufgestellt worden ist.

Damit ist bewiesen worden, daß alle Punkte der von der Wahl der ϱ-Zuordnung unabhängigen Menge M_ϱ auf einer Seite von $\mathfrak{E}$ liegt.

Wir zählen eine Kurve $\mathfrak{L}$ der Klasse $\mathfrak{A}$

1. *zu der Klasse $\mathfrak{A}_p$, wenn entweder die ganze Kurve $\mathfrak{L}$ oder die entsprechende Menge M_ϱ an positiver Seite von $\mathfrak{E}$ liegt,*

2. *zu der Klasse $\mathfrak{A}_n$, wenn eine von beiden an negativer Seite von $\mathfrak{E}$ liegt.*

e) Eigenschaft der Klassen $\mathfrak{A}_p$ und $\mathfrak{A}_n$. Wir werden die folgende Behauptung beweisen:

Gehört die Kurve $\mathfrak{L}$ zur Klasse $\mathfrak{A}_p$ (beziehungsweise $\mathfrak{A}_n$), so kann man eine solche positive Zahl σ finden, daß jede andere Vergleichskurve $\mathfrak{L}'$, die in der Umgebung $U_\sigma(\mathfrak{L})$ liegt, auch zur Klasse $\mathfrak{A}_p$ (beziehungsweise $\mathfrak{A}_n$) gehört.

Hat die Kurve $\mathfrak{L}$ mit $\mathfrak{E}$ keine gemeinsamen Punkte, so wird der Satz augenscheinlich (σ soll kleiner sein als die kleinste Entfernung $\mathfrak{L}$ von $\mathfrak{E}$).

Sei jetzt $\mathfrak{L}$ eine Kurve der Klasse $\mathfrak{A}$, die mit $\mathfrak{E}$ gemeinsame Punkte hat. Betrachten wir auf dem oben definierten Bogen $A_k \mathfrak{E} B_k$, für den die Ungleichheit $\bar{t}_k > t_k^* - 2\varepsilon$ gilt, noch zwei Punkte A und B, denen die Parameterwerte $t_k + \eta$ und $t_k - \eta$ entsprechen. Wählen wir η positiv und klein genug, daß noch

$$\bar{t}_k - \eta > (t_k + \eta)^* - 2\varepsilon. \tag{19}$$

Sei σ eine positive Zahl, die kleiner ist als die Entfernung des Bogens $A \mathfrak{E} B$ von der Kurve $\mathfrak{L}$. Sie ist auch kleiner als ϱ.

Betrachten wir eine Kurve $\mathfrak{L}'$, die in der Umgebung $U_\sigma(\mathfrak{L})$ liegt. Ich behaupte erstens, daß $\mathfrak{L}'$ zur Klasse $\mathfrak{A}$ gehört. In der Tat, hat die Umgebung $U_\sigma(\mathfrak{L})$ und infolgedessen auch die Kurve $\mathfrak{L}'$ keinen gemeinsamen Punkt mit dem Bogen $A \mathfrak{E} B$. Also, entweder hat $\mathfrak{L}'$ überhaupt keine gemeinsamen Punkte mit $\mathfrak{E}$, oder es gibt einen Bogen der Extremale $\mathfrak{E}$, der den Bogen $A \mathfrak{E} B$ enthält und der im Innern keine Punkte von $\mathfrak{L}'$ hat; nach (19) gilt für diesen Bogen die für die Klasse $\mathfrak{A}$ charakteristische Ungleichheit. In beiden Fällen gehört $\mathfrak{L}'$ zur Klasse $\mathfrak{A}$, was zu beweisen war.

Jetzt wollen wir beweisen, daß die Kurven $\mathfrak{L}$ und $\mathfrak{L}'$ zu derselben Klasse $\mathfrak{A}_p$ oder $\mathfrak{A}_n$ gehören. Dazu verändern wir etwas die Beweisführung

unter d). Wie dort, schließen wir aus (19) nach (14), daß

$$\bar{t}_k - \eta > t_k + \eta + 2\varepsilon. \tag{20}$$

Wir verschieben die Punkte C_k und E_k um η nach rechts, die Punkte D_k und F_k um η nach links. So erhalten wir Punkte mit entsprechenden Parameterwerten: A mit $t = t_k + \eta$, C mit $t = t_k + \frac{\varepsilon}{3} + \eta$, E mit $t = t_k + \frac{2\varepsilon}{3} + \eta$, F mit $t = \bar{t}_k - \frac{2\varepsilon}{3} - \eta$, D mit $t = \bar{t}_k - \frac{\varepsilon}{3} - \eta$, B mit $t = \bar{t}_k - \eta$. Der Ungleichheit (20) halber folgt auch jetzt F dem Punkte E nach. Wir führen die Bezeichnung $V = U_\varrho(C\mathfrak{E}D)$, $W = U_\varrho(E\mathfrak{E}F)$ ein (in der Fig. 4 sind diese Umgebungen mit punktierter Linie gezeichnet).

Setzen wir erstens voraus, daß $\mathfrak{L}'$ mit $\mathfrak{E}$ gemeinsame Punkte hat. Bilden wir für $\mathfrak{L}'$, wie unter d), die Umgebungen $V_{\mathfrak{L}'}$ und $W_{\mathfrak{L}'}$. Dann $V \prec V_{\mathfrak{L}}$, $V \prec V_{\mathfrak{L}'}$, $W \prec W_{\mathfrak{L}}$, $W \prec W_{\mathfrak{L}'}$. Man kann die Beweisführung unter d) wiederholen, indem man V und W anstatt $V_{\mathfrak{L}}$ und $W_{\mathfrak{L}}$ benutzt. So beweist man, daß jeder der Bogen $M\mathfrak{L}N$ und $M'\mathfrak{L}'N'$, die auf den Kurven $\mathfrak{L}$ und $\mathfrak{L}'$ dem Bogen $C\mathfrak{E}D$ bei einer beliebigen ϱ-Zuordnung entsprechen, ganz in V an einer Seite von $\mathfrak{E}$ liegt. Weiter beweist man, daß die Mengen M und M' der Punkte von $\mathfrak{L}$ und $\mathfrak{L}'$, die in V liegen, nicht leer sind und daß M dem Bogen $M\mathfrak{L}N$ und M' dem Bogen $M'\mathfrak{L}'N'$ angehört. Man bemerkt, daß $M \prec M_{\mathfrak{L}}$, $M' \prec M_{\mathfrak{L}'}$.

Sei Q' ein Punkt von M' und Q ein Punkt von $\mathfrak{L}$, dessen Entfernung von Q' kleiner als σ ist ($\mathfrak{L}'$ gehört der Umgebung $U_\sigma(\mathfrak{L})$ an). Die Entfernung des Punktes Q von $E\mathfrak{E}F$ ist kleiner als $\varrho + \sigma$, also auch als 2ϱ. Wie unter d) beweist man, daß Q nur den Punkten von $C\mathfrak{E}D$ entsprechen kann und daß er dem Bogen $M\mathfrak{L}N$ angehören muß. So haben wir zwei Punkte Q und Q' gebildet, von denen Q zu $M\mathfrak{L}N$ und Q' zu $M'\mathfrak{L}'N'$ gehört und deren Entfernung kleiner als σ ist.

Ich behaupte, daß Q und Q auf derselben Seite von $\mathfrak{E}$ liegen. In der Tat, wie unter d) beweist man, daß die Entfernung des Punktes Q von $B\mathfrak{E}A$ größer als ϱ, also auch als σ ist. Außerdem ist die Entfernung jedes Punktes der Kurve $\mathfrak{L}$, also auch des Punktes Q von dem Bogen $A\mathfrak{E}B$ größer als σ. Infolgedessen ist die Entfernung des Punktes Q von $\mathfrak{E}$ größer als σ. Also muß der Punkt Q', dessen Entfernung von Q kleiner als σ ist, auf derselben Seite von $\mathfrak{E}$ liegen wie Q.

Liegen Q und Q' auf derselben Seite von $\mathfrak{E}$, so trifft dasselbe für die Bogen $M\mathfrak{L}N$ und $M'\mathfrak{L}'N'$ zu, dann für ihre Teilmengen M und M', endlich für die Mengen $M_{\mathfrak{L}}$ und $M_{\mathfrak{L}'}$, von denen man weiß, daß jede auf einer Seite von $\mathfrak{E}$ liegt und daß sie beziehungsweise M und M' enthalten.

Hat $\mathfrak{L}'$ gemeinsame Punkte mit $\mathfrak{E}$, so haben wir das Geforderte bewiesen. Liegt $\mathfrak{L}'$ auf einer Seite von $\mathfrak{E}$, so genügt es, in dem Beweise die ganze Kurve $\mathfrak{L}'$ anstatt $M_{\mathfrak{L}'}$ zu setzen.

f) Eigenschaft der Klasse $\mathfrak{B}$. Jetzt wenden wir uns zur Betrachtung der Kurven $\mathfrak{L}$ von der Klasse $\mathfrak{B}$ zu. Wir werden beweisen, daß *überhaupt keine Kurven $\mathfrak{L}$ zur Klasse $\mathfrak{B}$ gehören.*

Setzen wir im Gegenteil voraus, daß es eine Kurve $\mathfrak{L}$ der Klasse $\mathfrak{B}$ gibt. Sei P_1 ein beliebiger gemeinsamer Punkt von $\mathfrak{E}$ und $\mathfrak{L}$, dem der Parameterwert t_1 entspricht. In dem Intervalle $t_1 \leqq t \leqq t_1^* - 2\varepsilon$ müssen noch Punkte von $\mathfrak{L}$ liegen. Sei t_2 der größte Parameterwert, der einem solchen Punkte P_2 entspricht. Dann

$$t_2 \leqq t_1^* - 2\varepsilon. \tag{21}$$

Betrachten wir jetzt zwei Fälle:

1. Der Punkt P_1 liegt in dem Intervalle $t_2 \leqq t \leqq t_2^* - 2\varepsilon$. Mit anderen Worten, $t_1 + d \leqq t_2^* - 2\varepsilon$. In diesem Falle hat $\mathfrak{L}$ mit $\mathfrak{E}$ zwei gemeinsame Punkte P_1 und P_2, von denen der Punkt P_2 dem Punkte P_1^*, und P_1 dem Punkte P_2^* wenigstens um 2ε vorangeht.

2. Der Punkt P_1 liegt nicht in dem Intervalle $t_2 \leqq t \leqq t_2^* - 2\varepsilon$. Jetzt ist $t_1 + d > t_2^* - 2\varepsilon$. In dem Intervalle $t_2 \leqq t \leqq t_2^* - 2\varepsilon$ müssen nach der Definition der Klasse $\mathfrak{B}$ Punkte von $\mathfrak{L}$ liegen. Sei t_3 der größte Parameterwert, der einem solchen Punkte P_3 entspricht. Da t_2 der größte Wert war, der die Ungleichheit (21) erfüllt, so hat man

$$t_3 > t_1^* - 2\varepsilon.$$

Daraus folgt nach dem Punkte 3. der Definition von ε, daß

$$t_3^* > t_1^{**} - \delta.$$

Nach (13) hat man $t_1^{**} > t_1 + d + 2\delta$ und daraus

$$t_3^* > t_1 + d + \delta.$$

Endlich folgt aus dem Punkte 1. der Definition von ε

$$t_1 + d < t_3^* - 2\varepsilon.$$

Also hat in diesem Falle $\mathfrak{L}$ mit $\mathfrak{E}$ drei gemeinsame Punkte P_1, P_2 und P_3, von denen P_2 dem Punkte P_1^*, P_3 dem Punkte P_2^*, und P_1 dem Punkte P_3^* wenigstens um 2ε vorangeht.

Aus der Beziehung zwischen den einander folgenden Punkten P_k folgt in beiden Fällen, daß der Komplementärbogen zu jedem der zwei oder drei Bogen, die auf $\mathfrak{E}$ gebildet worden sind, einen Teilbogen von Länge 2ε enthält. Im ersten Falle folgt daraus, daß jeder der Bogen $P_1 \mathfrak{E} P_2$ und $P_2 \mathfrak{E} P_1$ nicht kürzer als 2ε ist. Im zweiten Falle folgt, daß höchstens einer von den Bogen $P_1 \mathfrak{E} P_2$, $P_2 \mathfrak{E} P_3$ und $P_3 \mathfrak{E} P_1$ kürzer als ε sein kann (gäbe es zwei solche Bogen, so würde ihre Summe, die den Komplementärbogen zum dritten Bogen bildet, kürzer als 2ε sein).

Gibt es wirklich in dem zweiten Falle einen Bogen, der kürzer als ε ist, so schließen wir ihn dem vorangehenden Bogen an. Ist $P_1 \mathfrak{E} P_2$ (mit veränderter Bezeichnung) der zusammengesetzte Bogen, so geht P_2 dem Punkte P_1^* mehr als um ε voran.

Schließlich setzt sich $\mathfrak{E}$ in jedem Falle aus zwei oder drei Bogen zusammen, die nicht kürzer als ε sind. Sind P_i und P_k der Anfangs- und Endpunkt eines der Bogen, so geht P_k dem Punkte P_i^* mehr als um ε voran.

Betrachten wir eine beliebige ϱ-Zuordnung zwischen $\mathfrak{L}$ und $\mathfrak{E}$. Den Punkten P_k ($k = 1, 2$ oder $k = 1, 2, 3$ je nachdem wieviel Bogen es gibt), als Punkten von $\mathfrak{L}$, entsprechen gewisse Punkte von $\mathfrak{E}$. Sei S_k der Punkt oder, wenn es mehrere gibt, einer der Punkte, die P_k entsprechen. Aus dem Punkte 2. der Definition von ϱ folgt, daß S_k auf dem Bogen $t_k - \frac{\varepsilon}{3} < t < t_k + \frac{\varepsilon}{3}$ (wo t_k dem Punkte P_k auf $\mathfrak{E}$ entspricht) liegt. Da die Bogen $P_i \mathfrak{E} P_k$ nicht kürzer als ε sind, so haben die Punkte S_k dieselbe Ordnung auf $\mathfrak{E}$ wie die Punkte P_k.

Den zwei oder drei Bogen $S_i \mathfrak{E} S_k$ entsprechen in der ϱ-Zuordnung zwei oder drei Bogen $P_i \mathfrak{L} P_k$ der Kurve $\mathfrak{L}$, die zusammengenommen die ganze Kurve $\mathfrak{L}$ bilden. Jeder Bogen $P_i \mathfrak{L} P_k$ liegt in der Umgebung $U_\varrho(S_i \mathfrak{E} S_k)$. Bezeichnen wir mit P_i^- den Punkt, der dem Punkte P_i auf $\mathfrak{E}$ um $\frac{\varepsilon}{3}$ vorangeht, und mit P_k^+ den, der dem Punkte P_k um $\frac{\varepsilon}{3}$ nachfolgt, so ist $S_i \mathfrak{E} S_k$ ein Teilbogen von $P_i^- \mathfrak{E} P_k^+$. Infolgedessen liegt $P_i \mathfrak{L} P_k$ auch in der Umgebung $U_\varrho(P_i^- \mathfrak{E} P_k^+)$.

Erinnert man sich an den Punkt 1. der Definition von ϱ und macht von dem Hilfssatze des § 8 Gebrauch, so schließt man, daß der Bogen $P_i \mathfrak{E} P_k$ in $U_\varrho(P_i^- \mathfrak{E} P_k^+)$ ein Minimum liefert. Daher haben wir

$$J(P_i \mathfrak{E} P_k) \leqq J(P_i \mathfrak{L} P_k).$$

Summiert man entsprechende Ungleichheiten für alle zwei oder drei Bogen $P_i \mathfrak{E} P_k$, so bekommt man endlich

$$J(\mathfrak{E}) \leqq J(\mathfrak{L}),$$

was nach der Definition der Kurven $\mathfrak{L}$ (unter c) zum Widerspruch führt.

g) Beweis des Satzes. Seien zwei beliebige Kurven $\mathfrak{L}$: eine, $\mathfrak{L}_p$, liege auf der positiven, die zweite, $\mathfrak{L}_n$, auf der negativen Seite von $\mathfrak{E}$. *Setzen wir voraus, daß es eine Schar von Kurven $\mathfrak{L}_\alpha$ gibt, die stetig (gleichmäßig in bezug auf τ) von einem Parameter α im Intervalle $a \leqq \alpha \leqq b$ abhängen und die für $\alpha = a$ und $\alpha = b$ beziehungsweise mit $\mathfrak{L}_p$ und $\mathfrak{L}_n$ zusammenfallen.*

Gehört eine Kurve $\mathfrak{L}_{\alpha_0}$ zur Klasse $\mathfrak{A}_p$ (beziehungsweise $\mathfrak{A}_n$), so kann man nach e) eine Zahl σ finden, so daß jede Kurve $\mathfrak{L}$, die in $U_\sigma(\mathfrak{L}_{\alpha_0})$ liegt, noch zur Klasse $\mathfrak{A}_p$ (beziehungsweise $\mathfrak{A}_n$) gehört. Da die Kurven $\mathfrak{L}_\alpha$ stetig (gleichmäßig in bezug auf τ) von α abhängen, kann man eine positive Zahl γ finden, daß für $\alpha_0 - \gamma < \alpha < \alpha_0 + \gamma$ die Kurven $\mathfrak{L}_\alpha$ in $U_\sigma(\mathfrak{L}_{\alpha_0})$ liegen, also zur Klasse $\mathfrak{A}_p$ (beziehungsweise $\mathfrak{A}_n$) gehören.

Nach f) gehört keine Kurve $\mathfrak{L}_\alpha$ zur Klasse $\mathfrak{B}$.

Sei α_0 die untere Grenze der α-Werte, für welche die Kurve $\mathfrak{L}_\alpha$ zur Klasse $\mathfrak{A}_n$ gehört.

Gehört $\mathfrak{L}_{\alpha_0}$ zur Klasse $\mathfrak{A}_n$, so ist $\alpha_0 > a$ (weil $\mathfrak{L}_a = \mathfrak{L}_p$ zur Klasse $\mathfrak{A}_p$ gehört), und noch für die Werte $\alpha_0 - \gamma < \alpha \leqq \alpha_0$ gehört $\mathfrak{L}_\alpha$ zur Klasse $\mathfrak{A}_n$, was der Definition der unteren Grenze widerspricht.

Gehört $\mathfrak{L}_{\alpha_0}$ zur Klasse $\mathfrak{A}_p$, so ist $\alpha_0 < b$ (weil $\mathfrak{L}_b = \mathfrak{L}_n$ zur Klasse $\mathfrak{A}_n$ gehört), und noch für alle Werte $\alpha_0 \leqq \alpha < \alpha_0 + \gamma$ gehört $\mathfrak{L}_\alpha$ zur Klasse $\mathfrak{A}_p$, was auch zum Widerspruch führt.

Da $\mathfrak{L}_{\alpha_0}$ zu keiner Klasse gehören kann, so ist die vorausgesetzte Überführung unmöglich. Die Minimaxeigenschaft der Extremale $\mathfrak{E}$ ist bewiesen worden.

10. Notwendige Bedingung.

Satz. *Enthält die geschlossene Extremale $\mathfrak{E}$ den zweiten konjugierten Punkt eines ihrer Punkte und fallen die beiden Punkte nicht zusammen, so besitzt sie nicht die Minimaxeigenschaft.*

Mit anderen Worten, setzen wir voraus, daß mindestens für ein t

$$t^{(2)} < t + d. \tag{22}$$

Beweis. Sei A ein Punkt, dem der Wert t entspricht, für den (22) gilt (Fig. 5). Da $A^{(2)}$ dem Punkte A vorangeht, so kann man noch einen Punkt D finden, der dem Punkte $A^{(1)}$ folgt und dessen erster konjugierter Punkt $D^{(1)}$ dem Punkte A vorangeht. Wählen wir noch einen Punkt C zwischen $A^{(1)}$ und D und einen Punkt B zwischen $D^{(1)}$ und A.

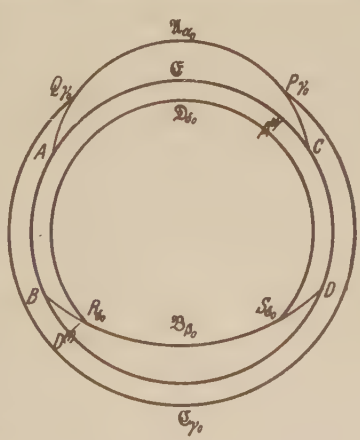


Fig. 5.

In *beliebiger* Umgebung U von $\mathfrak{E}$ können wir eine Schar $\mathfrak{A}_\alpha$ der Kurven einführen, die die Punkte A und C verbinden und die ganz analog zu denen von § 5 gebildet werden. Dasselbe betrifft die Schar $\mathfrak{B}_\beta$ der Kurven, die die Punkte D und B verbinden.

Wir verändern ein wenig die Definitionen der Scharen $\mathfrak{C}_\gamma$ und $\mathfrak{D}_\delta$. Die Kurven der Schar

$$\mathfrak{C}_\gamma: \qquad x = x_3(\tau, \gamma),\ y = y_3(\tau, \gamma),\ 0 \leqq \tau \leqq \tau_3(\gamma)$$

sollen folgende Eigenschaften besitzen:

1. Für jedes γ eines Intervalles $0 \leqq \gamma \leqq \gamma_0$ verbindet die Kurve $\mathfrak{C}_\gamma$ einen Punkt P_γ der Kurve $\mathfrak{A}_{\alpha_0}$, dem der Parameterwert $\tau = \tau_1(\alpha_0) - \gamma$ entspricht, mit einem Punkte Q_γ derselben Kurve $\mathfrak{A}_{\alpha_0}$, dem der Parameterwert $\tau = \gamma$ entspricht.

2. Für $\gamma = 0$ fällt die Kurve $\mathfrak{C}_\gamma = \mathfrak{C}_0$ mit $C\mathfrak{E}A$ zusammen.

3. und 4. bleiben mit den entsprechenden Nummern von § 5 gleichlautend.

Ähnlich sollen die Kurven der Schar

$$\mathfrak{D}_\delta: \qquad x = x_4(\tau, \delta),\ y = y_4(\tau, \delta),\ 0 \leqq \tau \leqq \tau_4(\delta)$$

folgende Eigenschaften besitzen:

1. Für jedes δ eines Intervalles $0 \leqq \delta \leqq \delta_0$ verbindet die Kurve $\mathfrak{D}_\delta$ einen Punkt R_δ der Kurve $\mathfrak{B}_{\beta_0}$, dem der Parameterwert $\tau = \tau_2(\beta_0) - \delta$ entspricht, mit dem Punkte S_δ derselben Kurve $\mathfrak{B}_{\beta_0}$, dem der Parameterwert $\tau = \gamma$ entspricht.

2. Für $\delta = 0$ fällt die Kurve $\mathfrak{D}_\delta = \mathfrak{D}_0$ mit $B\mathfrak{E}D$ zusammen.

3. und 4. sind mit den entsprechenden Nummern von § 5 gleichlautend.

Das Integral J ist längs der Kurve $A\mathfrak{A}_{\alpha_0}C\mathfrak{E}A$ kleiner als $J(\mathfrak{E})$. Wir wählen γ_0 hinreichend klein, daß J noch längs der Kurven $Q_\gamma \mathfrak{A}_{\alpha_0} P_\gamma \mathfrak{C}_\gamma Q_\gamma$ für $0 \leqq \gamma \leqq \gamma_0$ kleiner als $J(\mathfrak{E})$ sei. Ebenso wählen wir δ_0 so klein, daß J längs der Kurven $S_\delta \mathfrak{B}_{\beta_0} R_\delta \mathfrak{D}_\delta S_\delta$ kleiner als $J(\mathfrak{E})$.

Bezeichnen wir die geschlossenen Kurven $Q_{\gamma_0} \mathfrak{A}_{\alpha_0} P_{\gamma_0} \mathfrak{C}_{\gamma_0} Q_{\gamma_0}$ und $S_{\delta_0} \mathfrak{B}_{\beta_0} R_{\delta_0} \mathfrak{D}_{\delta_0} S_{\delta_0}$ mit $\mathfrak{L}_p$ und $\mathfrak{L}_n$. Dann gelten folgende Behauptungen:

1. Die Kurve $\mathfrak{L}_p$ liegt in U auf der positiven, $\mathfrak{L}_n$ auf der negativen Seite von $\mathfrak{E}$.

2. Das Integral J nimmt längs $\mathfrak{L}_p$ und $\mathfrak{L}_n$ einen kleineren Wert an als $J(\mathfrak{E})$.

3. Man kann $\mathfrak{L}_p$ in $\mathfrak{L}_n$ stetig in U so überführen, daß längs aller Übergangskurven J kleiner als $J(\mathfrak{E})$ ist.

Wie im § 5 zeigen wir nur, wie man den in 3. erwähnten Übergang vollenden kann. Zu diesem Zweck:

a) Betrachten wir die Kurve $Q_\gamma \mathfrak{A}_{\alpha_0} P_\gamma \mathfrak{C}_\gamma Q_\gamma$ und verändern γ von γ_0 bis 0. Für $\gamma = \gamma_0$ fällt diese Kurve mit $\mathfrak{L}_p$ zusammen, für $\gamma = 0$ mit $A\mathfrak{A}_{\alpha_0}C\mathfrak{E}A$.

b) Jetzt betrachten wir die Kurve $A\,\mathfrak{A}_{\alpha_0}\,C\,\mathfrak{E}\,D\,\mathfrak{B}_\beta\,B\,\mathfrak{E}\,A$ und verändern β von 0 bis β_0. Für $\beta = 0$ fällt diese Kurve mit der Endkurve von a) zusammen, für $\beta = \beta_0$ mit $A\,\mathfrak{A}_{\alpha_0}\,C\,\mathfrak{E}\,\mathfrak{B}_{\beta_0}\,B\,\mathfrak{E}\,A$.

c) Jetzt betrachten wir die Kurve $A\,\mathfrak{A}_\alpha\,C\,\mathfrak{E}\,D\,\mathfrak{B}_{\beta_0}\,B\,\mathfrak{E}\,A$ und verändern α von α_0 bis 0. Für $\alpha = \alpha_0$ fällt diese Kurve mit der Endkurve von b) zusammen, für $\alpha = 0$ mit $A\,\mathfrak{E}\,D\,\mathfrak{B}_{\beta_0}\,B\,\mathfrak{E}\,A = B\,\mathfrak{E}\,D\,\mathfrak{B}_{\beta_0}\,B$.

d) Endlich betrachten wir die Kurve $R_\delta\,\mathfrak{D}_\delta\,S_\delta\,\mathfrak{B}_{\beta_0}\,R_\delta$ und verändern δ von 0 bis δ_0. Für $\delta = 0$ fällt diese Kurve mit der Endkurve von c) zusammen, für $\delta = \delta_0$ mit $\mathfrak{L}_n$.

Man kann leicht bestätigen, daß dieser Übergang alle Forderungen erfüllt. Man sieht auch, daß es für beliebig kleines ϱ alle Konstruktionen so durchzuführen möglich ist, daß alle Übergangskurven der Umgebung $U_\varrho(\mathfrak{E})$ ordentlich angehören.

Also besitzt die Extremale $\mathfrak{E}$ die Minimaxeigenschaft nicht.

11. Bemerkung.

Es ist der Fall geblieben, in welchem (22) für kein t und (12) nicht für alle t gilt. Mit anderen Worten, in welchem für alle t

$$t^{(2)} \geqq t + d$$

und mindestens für ein t das Gleichheitszeichen gilt.

Dieser Grenzfall kann nicht ohne Bedingungen höchster Ordnung entschieden werden.

12. Eine Bemerkung über Voraussetzungen.

Bei der Definition der Minimaxeigenschaft im Falle der geschlossenen Extremalen haben wir eine Einschränkung gemacht, daß die Vergleichskurven der Umgebung $U_\varrho(\mathfrak{E})$ *ordentlich* angehören. Man könnte diese Voraussetzung durch andere ersetzen, zum Beispiel, daß die Vergleichskurve in $\mathfrak{E}$ stetig überführbar ist, oder daß sie denselben Umlaufssinn wie $\mathfrak{E}$ hat. Im letzten Falle muß man den Umlaufssinn streng definieren.

Wir werden an einem Beispiele zeigen, daß man sich nicht ohne jede solche Voraussetzung behelfen kann.

Sei

$$J = \int \frac{2\sqrt{(x^2+y^2)\,(x\dot{x}+y\dot{y})^2 + (x\dot{y}-y\dot{x})^2} + [1-k^2(\sqrt{x^2+y^2}-1)^2]\,(x\dot{y}-y\dot{x})}{x^2+y^2}\,dt$$

oder

$$J = \int \{2\sqrt{\dot{r}^2+\dot{\theta}^2} + [1-k^2(r-1)^2]\,\dot{\theta}\}\,dt,$$

wo r, θ Polarkoordinaten sind, die Punkte die Ableitungen nach t bezeichnen, und $0 < k < 1$.

Man sieht leicht, daß die Kurve $\mathfrak{E}$

$$x = \cos t,\ y = \sin t$$

eine Extremale ist, daß längs $\mathfrak{E}$ die Bedingungen (3), (4), (5) erfüllt sind Zum Beweis ist am bequemsten sich, der Polarkoordinaten zu bedienen Endlich berechnet man, daß für jedes t

$$t^{(1)} - t = \frac{\pi}{k},$$

$$t^{(2)} - t = \frac{2\pi}{k},$$

daß, mit anderen Worten, die Voraussetzungen des Satzes des § 9 erfüllt sind

Trotzdem gilt die Minimaxeigenschaft nicht. In der Tat betrachten wir die Schar der Kurven $\mathfrak{L}_\alpha$:

$$x = (1+\alpha)\cos t,\ y = -(1-\alpha)\sin t,$$

wo $|\alpha| < 1$. Das Integral längs der Kurve $\mathfrak{L}_\alpha$ ist gleich

$$J(\mathfrak{L}_\alpha) = 2\pi(1 + k^2\alpha^2) < 4\pi,$$

während

$$J(\mathfrak{E}) = 6\pi.$$

Wählt man eine beliebige Zahl ϱ, so genügt noch $|\omega| < \varrho$ zu setzen, um eine stetige Schar der Kurven $\mathfrak{L}_\alpha$ zu haben, die in $U_\varrho(\mathfrak{E})$ an beiden Seiten von $\mathfrak{E}$ liegen. Da

$$J(\mathfrak{L}_\alpha) < J(\mathfrak{E}),$$

so hat $\mathfrak{E}$ die Minimaxeigenschaft nicht.

Man bemerkt, daß die Kurven $\mathfrak{L}_\alpha$ der Umgebung $U_\varrho(\mathfrak{E})$ nicht ordentlich angehören, was die Notwendigkeit einer Einschränkung rechtfertigt.

13. Geschlossene geodätische Linien auf einem Kreistorus.

Betrachten wir einen Kreistorus, d. h. eine Fläche, die durch Umdrehung eines Kreises um eine in seiner Ebene liegende und ihn nicht schneidende Achse entsteht. Sei r der Radius des erzeugenden Kreises (des Meridianes) und R die größte Entfernung der Kreispunkte von der Achse (Äquatorradius).

Indem wir uns auf den Satz von Herrn Birkhoff[8]) über die geodätischen Linien auf einem topologischen Torus stützen, behaupten wir, daß der Kreistorus mindestens eine geschlossene geodätische Linie enthält, die die Minimaxeigenschaft besitzt und in den Kehlkreis (Minimum der Länge) stetig überführbar ist.

---

[8]) Fußnote [1]). Siehe auch Birkhoff, Dynamical Systems (1927), S. 133.

Dies kann nicht der Äquatorkreis sein. In der Tat, berechnet man, daß längs des Äquatorkreises (t bezeichnet die Länge)

$$t^{(1)} - t = \pi\sqrt{Rr},$$
$$t^{(2)} - t = 2\pi\sqrt{Rr},$$

woraus die Behauptung des Satzes von § 10 folgt (die Länge des Äquatorkreises ist $2\pi R$ gleich). Daher kann der Äquatorkreis nicht die Minimaxeigenschaft besitzen.

Was die Meridiankreise betrifft, so sind diese nicht in den Kehlkreis stetig überführbar.

Also haben wir bewiesen, daß es eine geschlossene geodätische Linie gibt, die stetig in den Kehlkreis überführbar ist, und die nicht mit dem Äquatorkreise zusammenfällt. Durch Umdrehung um die Torusachse entsteht aus ihr *eine einparametrige Schar von geschlossenen geodätischen Linien des Kreistorus, die in den Kehlkreis stetig überführbar sind.*

Die Existenz dieser Schar ist vom Verfasser auch mit der Minimaxmethode durch ein anderes Verfahren bewiesen worden[9]).

Warschau, den 21. November 1930.

---

[9]) Monatshefte für Mathematik und Physik **38** (1931), S. 53.

(Eingegangen am 29. November 1930.)

Über einen neuen Zweig geometrischer Axiomatik und eine neue Art von analytischer Geometrie.

Von

G. Thomsen in Rostock.

Einleitung.

Wir stellen zuerst einige bekannte Tatsachen aus der geometrischen Grundlagenforschung zusammen und schließen an diese dann (auf Seite 673 ff.) einen Bericht über Inhalt und Ziel dieser Arbeit an.

Ein großer Teil der Hilbertschen Untersuchungen über die Grundlagen der Geometrie [1]) besteht in der Aufdeckung der tieferen logischen Zusammenhänge jener Verwandtschaft von Elementargeometrie und Körperalgebra auf der Descartes' analytische Geometrie beruht. Besonders weit ist die Analyse auf dem Teilgebiet der Geometrie getrieben worden, das man als „Geometrie der reinen Schnittpunktssätze" zu bezeichnen pflegt. Wenn wir uns auf die Ebene beschränken, so handelt diese Teilgeometrie ausschließlich von solchen geometrischen Sätzen, die reine Koinzidenzbeziehungen von Punkten und Geraden ausdrücken, ohne Aussagen über Anordnung, Stetigkeit oder Kongruenz zu machen. Die beiden Haupttatsachen sind hier die folgenden:

I. Nimmt man einen beliebigen Körper, so kann man sich eine zugehörige „ebene analytische Geometrie" konstruieren, indem man erstens ein geordnetes Paar (x, y) von Körperelementen als *Punkt* interpretiert, wobei x und y die Koordinaten des Punktes heißen; indem man weiter die Menge aller derjenigen Punkte als *Gerade* auffaßt, deren Koordinatenelemente x, y dieselbe lineare Gleichung $ux + vy + w = 0$ mit festen Körperelementen u, v und w erfüllen, wobei u und v nicht gleichzeitig das Nullelement sein dürfen. Hierbei bezeichnet man die Elemente u, v und w als Koordinaten der Geraden und hat nur zu beachten, daß zwei Geraden deren beiden Koordinatentripel sich nur um einen gemeinsamen Faktor

[1]) Vgl. D. Hilbert, Grundlagen der Geometrie, 7. Auflage 1930 (Teubner), im besonderen Kap. V und VI.

nterscheiden, identisch sind. Nun konstruiert man sich die „Geometrie", ndem man drittens das Zugehören eines Punktes zu einer solchen (durch ine lineare Gleichung definierten) Punktmenge als *„vereinigte Lage von Punkt und Gerade"* definiert. Aus den Körperaxiomen ergibt sich jetzt, laß zwei Gerade (u, v, w) und (u', v', w') dann und nur dann keinen Punkt gemein haben, wenn

$$1) \qquad uv' - u'v = 0$$

gilt; solche Gerade heißen *parallel. In dieser abstrakten Geometrie ist,* wie sich jetzt zeigt — von gewissen Existenzaussagen abgesehen —, *jeder eine Schnittpunktssatz richtig, der in der gewöhnlichen euklidischen Geometrie gilt.* Von besonderem Reiz ist es hier bekanntlich, Geometrien mit ndlich vielen Punkten und Geraden zu konstruieren, in denen jeder dieser chnittpunktssätze gültig ist[2]).

II. Die zweite tiefere Erkenntnis, die auf Hilbert und Hessenberg urückgeht, ist in gewissem Sinne eine Umkehrung des vorigen: Es sind wei Systeme von Dingen vorgelegt, Punkte und Gerade, wobei gewisse Punkte mit gewissen Geraden durch eine Beziehung „vereinigte Lage" oder „Koinzidenz" verknüpft sind. Für diese Beziehung gelten die vier Axiome: . Es gibt drei nicht mit einer und derselben Geraden koinzidierende Punkte. 2. Zu zwei Punkten gibt es immer genau eine mit beiden gleichzeitig koinzidierende Gerade. 3. Zu einem Punkt P und einer mit ihm nicht koinzidierenden Geraden g gibt es immer genau eine Gerade h, welche mit P in vereinigter Lage ist, ohne daß es zu g und h einen mit beiden Geraden gleichzeitig konzidierenden Punkt gibt (Parallelenaxiom). . Es gilt für die Koinzidenzbeziehung von Punkten und Geraden der ogenannte *Schnittpunktssatz von Pascal*[3]). *Eine jede solche, in abstracto d. h. bis auf Isomorphismen) gegebene Geometrie läßt sich nun nach Hilbert und Hessenberg als analytische Geometrie eines Körpers in dem nter* I. *angegebenen Sinne gewinnen*[4]).

Über den Rahmen der reinen Schnittpunktsgeometrie hinausgehend, at man auf diese Verwandtschaft zwischen Geometrie und Körperalgebra uch eine logische Analyse der Kongruenz-, Anordnungs- und Stetigkeitsegriffe der Geometrie gegründet. In Analogie zur gewöhnlichen analyschen Geometrie im reellen Körper kann man in einer allgemeinen Körper-

[2]) Nimmt man den Körper der Restklassen modulo 5, so erhält man eine Geoetrie mit 25 Punkten und 30 Geraden.

[3]) Vgl. Hilbert, a. a. O. Kap. VI.

[4]) Am weitesten ist die logische Analyse der Schnittpunktsätze durchgeführt dem Buch von Reidemeister, Grundlagen der Geometrie. Julius Springer, (Berlin, 930). Grundlehren der math. Wissensch., Band XXXII.

geometrie als *Entfernungsquadrat* zweier Punkte (x, y) und (x', y') das Körperelement

$$Q = (x - x')^2 + (y - y')^2 \tag{2}$$

einführen. Nennen wir eine Gerade (u, v, w) mit $u^2 + v^2 = 0$ (ohne daß $u = v = 0$ ist) *isotrop* und eine solche mit $u^2 + v^2 \neq 0$ *anisotrop*, so kann man als „*Kosinusquadrat des Winkels zweier anisotroper Geraden*“ (u, v, w) und (u', v', w') das Element

$$C = \frac{(uu' + vv')^2}{(u^2 + v^2)(u'^2 + v'^2)} \tag{3}$$

definieren. Die durch

$$x = 2a - x^*; \quad y = 2b - y^* \tag{4}$$

vermittelte eineindeutige Abbildung $(x, y) \rightarrow (x^*, y^*)$ aller Punkte der Geometrie bezeichnen wir als *Spiegelung am Punkte* (a, b). Die „*Spiegelung an der anisotropen Geraden* (u, v, w)“ definieren als die Abbildung

$$x = x^* - 2u\frac{ux^* + vy^* + w}{u^2 + v^2}; \quad y = y^* - 2v\frac{ux^* + vy^* + w}{u^2 + v^2}. \tag{5}$$

In Übereinstimmung mit (3) wird man weiter zwei Gerade als *senkrecht* bezeichnen, wenn

$$uu' + vv' = 0 \tag{6}$$

gilt. Dabei können wir in (6) auch isotrope Gerade zulassen. Der Satz, daß zwei senkrechte Gerade immer einen Schnittpunkt haben, also nie parallel sind, gilt nach (1), (6) in allgemeinen Körpern nur für anisotrope Geradenpaare. Denn wenn für die Geraden (u, v, w) und (u', v', w') gilt: $u = u'$, $v = v'$, $w \neq w'$ und $u^2 + v^2 = 0$, ohne daß u und v beide das Nullelement sind, so sind die Gleichungen

$$uv' - vu' = 0, \quad uu' + vv' = 0 \tag{7}$$

beide erfüllt. Die beiden voneinander verschiedenen Geraden sind also sowohl senkrecht wie parallel. Wir schließen nun die Existenz isotroper Geraden aus und sichern damit die Existenz des Schnittpunktes zweier beliebiger senkrechter Geraden unserer Geometrie, indem wir verlangen, daß $u^2 + v^2 = 0$ nur für $u = v = 0$ möglich ist, oder, was auf dasselbe hinauskommt, indem wir an unsern Körper die Forderung stellen:

Axiom A. *Die Gleichung* $x^2 = -1$ *ist für kein Element* x *erfüllbar.*

Aus A folgt, daß unser Körper nicht die Charakteristik 2 besitzen kann, denn in einem solchen Körper würde die Einheit mit dem Element -1 zusammenfallen und selbst eine Lösung von $x^2 = -1$ liefern. Man kann in Körpern einer von 2 verschiedenen Charakteristik jetzt den *Mittelpunkt* (m, n) eines Punktepaares (x, y), (x', y') definieren:

$$m = \frac{x + x'}{2}; \quad n = \frac{y + y'}{2}. \tag{8}$$

Ebenso läßt sich jetzt auch das *Mittellot* (die Symmetrielinie) eines Punktepaares erklären.

In einer Geometrie, die zu einem dem Axiom A genügenden Körper gehört, brauchen zwei Gerade keine Mittellinie (*Winkelhalbierende* oder Spiegelungsachse) zu besitzen. Denn wenn $(0, 1, 0)$ und $(1, q, 0)$ zwei vorgelegte Geraden sind, und wenn man die Spiegelungsachse (u, v, w) nach (5) zu bestimmen sucht, so transformiert sich die Gleichung der ersten der beiden Geraden $y = 0$ in die Form der Bildgeraden

$$(u^2 + v^2)\,y^* - 2v(ux^* + vy^* + w) = 0. \tag{9}$$

Soll diese Gerade mit $(1, q, 0)$ übereinstimmen, so muß gelten:

$$uv \neq 0;\ w = 0 \quad \text{und} \quad \left(\frac{u}{v} + q\right)^2 = q^2 + 1.$$

Soll $u : v$ bestimmbar sein, so muß also die Gleichung $x^2 = q^2 + 1$ mindestens eine Lösung für x besitzen. Um nun die Existenz mindestens einer Mittellinie je zweier Geraden sicherzustellen, fordern wir jetzt für unsern Körper:

Axiom B. *Die Gleichung $x^2 = q^2 + 1$ hat für jedes vorgegebene Körperelement q mindestens eine Lösung x.*

Wie leicht einzusehen, gibt es in einer Geometrie, die zu einem Körper mit den Eigenschaften A und B gehört, zu zwei nichtparallelen Geraden immer genau zwei Mittellinien, zu zwei parallelen Geraden [wie übrigens schon mit Axiom A allein folgt] genau eine. Auf Grund von Axiom B ist es jetzt auch möglich, auf einer mit einem daraufliegenden Punkt P gegebenen Geraden g genau zwei Punkte Q und $\bar{Q}$ zu finden, so daß die Entfernungsquadrate der Paare PQ und $P\bar{Q}$ dem Entfernungsquadrat eines gegebenen Paares MN gleich sind. Ebenso gibt es zu einer Geraden g und einem auf ihr liegenden Punkt P jetzt genau zwei Gerade h und $\bar{h}$ durch P, so daß das „Kosinusquadrat des Winkels“ von $\{g, h\}$ sowie von $\{g, \bar{h}\}$ gleich demjenigen eines beliebigen vorgegebenen nichtparallelen und nicht senkrechten Geradenpaares wird.

Man sieht nun ohne weiteres ein: *In einem Körper, in dem die Axiome* A *und* B *erfüllt sind, kann -1 nie gleich einer Quadratsumme beliebigvieler Elemente sein.*

Für Körper mit dieser letzteren Eigenschaft haben nun Artin und Schreier [5]) gezeigt, daß sie stets *geordnet* werden können in dem folgenden Sinne: Es ist möglich, die von der Null verschiedenen Elemente des Körpers so in zwei Klassen zu teilen, so daß erstens gilt: Wenn a zu einer der beiden Klassen gehört, so gehört $-a$ zur andern; und daß zweitens gilt: Wenn a und b zur ersten Klasse gehören, so gehören auch

[5]) Hamburger Abhandlungen **5** (1927), S. 85.

$a+b$ und ab zur ersten Klasse. Für Körper mit der ausgesprochenen Eigenschaft ist diese Klasseneinteilung im allgemeinen sogar auf mehrere Arten möglich. [Wir erhalten die gewöhnliche Ordnungsbeziehung „größer" ($>$) aus einer solchen Klasseneinteilung, wenn wir a immer dann $>b$ nennen, falls $a-b$ zur ersten Klasse gehört[6]).]

In einer Geometrie, die zu einem mit einer der möglichen Ordnungen versehenen Körper mit den Eigenschaften A und B gehört, können wir nun für die drei Punkte (x, y), (x', y'), (x'', y'') einer Geraden, die Relation *„zwischen"* auf folgende Weise erklären: Wir sagen, der Punkt (x', y') liegt zwischen den anderen beiden dann und nur dann, wenn weder die beiden Elemente $x-x'$ und $x'-x''$ zu verschiedenen Klassen gehören noch auch die beiden Elemente $y-y'$ und $y'-y''$ zu verschiedenen Klassen gehören. Wie unschwer einzusehen, gelten jetzt die elementaren Anordnungsaxiome einschließlich des Axioms von Pasch. *Überhaupt zeigt sich, daß in unserer Geometrie jetzt die sämtlichen ebenen Axiome der Hilbertschen Axiomgruppen* I—IV *erfüllt sind, also alle Axiome des zum Aufbau der ebenen Geometrie hinreichenden Hilbertschen Systems mit Ausnahme der Stetigkeitsaxiome.*

Es lassen sich in ihr alle Konstruktionen mit Lineal und Eichmaß ausführen, in ihr läßt sich der eigentliche Grundbestand an Sätzen der Elementargeometrie beweisen. Aus diesen Gründen wollen wir einen Körper, der den Axiomen A und B genügt, einen *„geometrischen Körper"* nennen. Mit der Hilbertschen Streckenrechnung läßt sich natürlich auch umgekehrt zeigen, daß sich jede Geometrie, die den Hilbertschen Axiomen I bis IV genügt, als analytische Geometrie eines geometrischen Körpers gewinnen läßt. Aus der Eigenschaft, daß -1 nie Quadratsumme sein kann, folgert man leicht, daß ein geometrischer Körper unendlichviele Elemente besitzen muß. Es gibt aber schon geometrische Körper mit abzählbar unendlichvielen Elementen.

In der Geometrie eines jeden geometrischen Körpers gibt es eine diesem zugeordnete *Gruppe kongruenter Abbildungen* $(x, y) \rightarrow (x^*, y^*)$

$$(10)\quad \begin{cases} \begin{cases} x = ax^* + by^* + r, \\ y = cx^* + dy^* + s, \end{cases} \\ \text{wobei für die sechs Gruppenparameter } a, b, c, d, r, s \text{ die Gleichungen bestehen müssen:} \\ a^2 + c^2 = b^2 + d^2 = 1; \qquad ab + cd = 0. \end{cases}$$

Diese Gruppe läßt sich aus den Spiegelungen an Geraden (5) erzeugen.

[6]) Es gelten dann die von Hilbert verwandten 4 Axiome für die $>$-Beziehung. Vgl. Grundlagen a. a. O. Kap. III, Seite 52, Axiom 13—16.

Diese Arbeit, zu welcher besonders Untersuchungen von Hjelmslev die Anregung gegeben haben. *strebt auf das Ziel hin, dem Descartes-Hilbertschen „Körperkalkül" eine ganz andere Art analytischer Geometrie gegenüberzustellen, die sich auf einen reinen „Gruppenkalkül" gründet.* Im folgenden werden gewisse erste Ansätze zu einer solchen rein gruppentheoretischen Geometrie gemacht. In Verbindung damit ergibt sich dann auch ein besonderer Zweig geometrischer Axiomatik.

Die Gruppe der kongruenten Abbildungen der Ebene, die der euklidischen Geometrie zugrunde liegt, läßt sich bekanntlich erzeugen aus den Spiegelungen an den Geraden (vgl. (5)) und den Spiegelungen an den Punkten (vgl. (4)). (Darauf, daß sogar die Spiegelungen an den Geraden allein schon die Gruppe erzeugen, wollen wir hier keinen Wert legen.) Diese Spiegelungen an den Geraden und Punkten sind bekanntlich außer der Identität die einzigen *involutorischen* kongruenten Abbildungen der Ebene, d. h. solche Abbildungen, die, zweimal hintereinander ausgeführt, die Identität ergeben. Bezeichnen wir die identische kongruente Abbildung mit 1, mit P und a aber die Spiegelungen am Punkte P bzw. an der Geraden a, so gelten also die Gruppenrelationen

$$PP=1; \quad aa=1. \tag{11}$$ [6a])

Es lassen sich nun eigentlich alle einfachen geometrischen Lagebeziehungen zwischen Punkten und Geraden kennzeichnen durch Gruppenrelationen zwischen den Spiegelungen an den betreffenden Elementen. Z. B. sind zwei voneinander verschiedene Gerade a und b dann und nur dann senkrecht, wenn die beiden Spiegelungen a, b an ihnen vertauschbar sind, d. h. wenn $ab=ba$ gilt, oder

$$abab=1. \tag{12}$$

Ein Punkt P liegt dann und nur dann auf einer Geraden a, wenn die Spiegelungen P und a vertauschbar sind, was man als

$$aPaP=1 \tag{13}$$

schreiben kann. Ebenso kann man zeigen (für alle diese Tatsachen finden sich die kurzen Beweise im § 2 zusammengestellt): Drei Gerade a, b und c gehören dann und nur dann demselben Büschel an (gehen durch denselben Punkt oder sind alle drei parallel), wenn

$$abcabc=1 \tag{14}$$

gilt. Der Punkt S ist Mittelpunkt der Strecke PQ, wenn

$$PSQS=1 \tag{15}$$

[6a]) Gruppenrelationen denken wir von links nach rechts gelesen.

gilt. Die Gerade a ist Mittellot der Strecke PQ, falls

$$PaQa = 1. \tag{16}$$

Die Gerade d ist Mittellinie (Winkelhalbierende) des Geradenpaares a, b, wenn

$$adbd = 1 \tag{17}$$

gilt. Wir wollen es hier mit dieser Auswahl bewenden lassen. (Ausführliches vgl. § 2.)

Natürlich kann man die beiden Arten von Spiegelungen, die an den Geraden und an den Punkten, leicht durch innere, allein aus der Struktur der Gruppe der kongruenten Abbildungen entspringende Eigenschaften voneinander unterscheiden. Das geht z. B. auf die folgende Weise: Wir prüfen ein herausgegriffenes (von der Identität verschiedenes) involutorisches Element β daraufhin, ob es sich mit mindestens einem weiteren involutorischen Element ϱ zu einem Produktelement $\beta\varrho$ zusammensetzen läßt, welches die Periode 3 besitzt, d. h. für welches $\beta\varrho\beta\varrho\beta\varrho = (\beta\varrho)^3 = 1$ gilt. Kann man zu einem Element β kein solches ϱ hinzufinden, dann haben wir eine Spiegelung an einem Punkt, im entgegengesetzten Falle aber eine solche an einer Geraden. In der Tat: Zu jeder Geraden β gibt es eine weitere ϱ, die mit ihr den Winkel 60^0 einschließt. Nun ist die Abbildung $\beta\varrho$, die durch Aneinanderreihen der Spiegelungen an β und an ϱ entsteht, aber eine Drehung um deren Schnittpunkt, und zwar um das Doppelte des Winkels, den sie einschließen, hier also um 120^0. Daher ist hier wirklich die dreifache Drehung $(\beta\varrho)^3$ die Identität. Ist β dagegen eine Punktspiegelung, so kann für keine Geradenspiegelung ϱ das Produkt $\beta\varrho\beta\varrho\beta\varrho = 1$ werden. Denn eine Punktspiegelung ist eine gleichsinnige, eine Geradenspiegelung aber eine ungleichsinnige kongruente Abbildung, und daher ist die Abbildung $(\beta\varrho)^3$, in der eine ungerade Anzahl von Geradenspiegelungen vorkommt, ebenfalls ungleichsinnig (eine Umlegung). Eine solche kann aber natürlich nie gleich der Identität sein. Ist ϱ aber ebenso wie β eine Punktspiegelung, so ist $\beta\varrho$ eine Translation um den doppelten Vektor der Strecke von β nach ϱ, die Abbildung $(\beta\varrho)^3$ ist also nie die Identität. Damit ist unsere Behauptung bewiesen[7]).

In ähnlicher Weise, wie wir oben zu jedem Körper eine analytische Geometrie konstruiert haben, können wir jetzt *zu irgendeiner vorgelegten abstrakten Gruppe eine zugehörige Geometrie konstruieren.* Wir teilen nämlich die von der Identität verschiedenen involutorischen Elemente in zwei Klassen, je nachdem ob solch ein Element β sich mit mindestens

[7]) Ein einfacheres Kriterium ist vielleicht noch das folgende: Eine Punktspiegelung läßt sich immer, eine Geradenspiegelung nie als Quadrat einer kongruenten Abbildung darstellen. Letztere kongruente Abbildung hat dann die Periode 4. Wir haben obiges Kriterium vorgezogen, weil es nur von Elementen der Periode 3 handelt.

einem andern involutorischen Element ϱ zu einem Produktelement von der Periode 3 zusammensetzen läßt oder nicht. Die Elemente der ersten Klasse nennen wir Gerade, die der zweiten Punkte. Wir bezeichnen dann zwei Gerade als senkrecht, wenn (12) gilt, wir nennen einen Punkt P auf der Geraden a liegend, wenn (13) gilt. Drei Gerade sollen einem Büschel angehörend heißen, wenn (14) gilt. In entsprechender Weise erklärt man dann nach (15), (16), (17) die Begriffe Mittelpunkt und Mittellot eines Punktepaares sowie Mittellinie eines Geradenpaares. Nach § 2 kann man die Zahl solcher ohne weiteres möglicher Definitionen noch vermehren. Natürlich werden nur Gruppen Interesse haben, bei denen die beiden angegebenen Klassen involutorischer Elemente nicht leer sind. Weiter wird für eine Gruppe und die Untergruppe, die durch deren involutorischen Elemente erzeugt wird, die Geometrie dieselbe sein. Nur auf letztere kommt es an.

Für jede solche Geometrie kann man nun fragen, welche der elementargeometrischen Sätze in ihr noch gültig sind und welche nicht. Gewisse Sätze allereinfachster Art gelten in allen diesen Geometrien. Z. B. folgt allein auf Grund der Tatsache, daß die Formeln (13) bis (17) Gruppenrelationen darstellen, rechnerisch der Satz: Eine Mittellinie des Geradenpaares a, b gehört mit a und b demselben Büschel an, nach (17), (14) in Formeln: Aus $adbd = 1$ folgt $abdabd = 1$

$$\{adbd = 1\} \rightarrow \{abdabd = 1\}. \tag{18}$$

Aus der Voraussetzung folgt nämlich, wenn man von links mit $d \cdot a$ und von rechts mit a multipliziert, unter Berücksichtigung von (11), daß $bda = d$ ist, und wegen $dd = 1$ folgt $bdabda = 1$, daraus ergibt sich dann durch Multiplizieren von links sowohl wie von rechts mit a das gewünschte $abdabd = 1$.

Es lassen sich aber lange nicht alle geometrischen Sätze durch reines Gruppenrechnen beweisen. So kann man z. B. leicht eine „Pseudo-Gruppengeometrie" aufstellen, in der folgender Satz nicht gilt:

$$\left\{\begin{array}{l} abcabc = 1 \\ abdabd = 1 \end{array}\right\} \rightarrow \{bcdbcd = 1\}. \tag{19}$$

In Worten: Es gilt nicht der Satz: Wenn die drei Geraden a, b und c einem Büschel angehören und ebenso die drei Geraden a, b und d, so gehören auch die drei Geraden a, c und d einem Büschel an. Dieser Satz ist somit allein auf Grund der Gruppenaxiome nicht zu beweisen. Dieses Beispiel rückt zugleich ins rechte Licht, daß die Gruppengeometrien noch sehr weit von der gewohnten Geometrie abweichen werden, wohl wesentlich mehr als die eingangs besprochenen Körpergeometrien. Z. B. braucht der Mittelpunkt eines Punktepaares nicht auf deren Mittellot zu liegen und

ebensowenig auf einer Geraden, die durch die beiden Punkte hindurchgeht. Für zwei Gerade, die sich im Punkte P schneiden, braucht eine Winkelhalbierende nicht gleichfalls durch P hindurch zu gehen. Zwei Punkte können mehrere, unendlichviele oder auch keine Verbindungsgerade besitzen usw.

Wie diese Beispiele zeigen, ist durch den gruppentheoretischen Standpunkt eine neue Schicht geometrischer Axiomatik angeschnitten. Von gewissem Reiz ist es auch hier, sich Geometrien mit endlichvielen Punkten und Geraden zu konstruieren, indem man von endlichen Gruppen ausgeht. Man erhält eine solche z. B., wenn man die vollständige Permutationsgruppe von vier Elementen nimmt. Nach dem Kriterium über die Produktelemente der Periode 3 zerfallen die von der Identität verschiedenen involutorischen Permutationen hier in zwei Klassen, in die Vertauschungen eines Paares unter Festhaltung der anderen beiden Elemente und in die Vertauschungen eines Paares unter gleichzeitiger Vertauschung der beiden übrigbleibenden Elemente, und zwar sind dabei die ersteren die Geraden, die letzteren die Punkte der Geometrie. Es gibt somit sechs Gerade und drei Punkte. In dieser Geometrie hat ein Punktepaar, wie sich zeigt, stets entweder kein Mittellot oder zwei Mittellote, aber nie eine Verbindungsgerade und nie einen Mittelpunkt. Ein Geradenpaar hat entweder keine oder nur eine Mittellinie. Es gilt der Satz von den drei Winkelhalbierenden eines Dreiecks in dem folgenden Sinne: Sind sechs Geraden vorgelegt a, b, c, p, q, r, so daß gemäß (17) gilt: $arbr = 1$; $bpcp = 1$ und $cqaq = 1$, so gehören die drei Geraden p, q und r notwendig einem Büschel im Sinne der Definition (14) an[7a]):

$$(20) \qquad \{arbr = 1;\ bpcp = 1;\ cqaq = 1\} \rightarrow \{pqrpqr = 1\}.$$

Es gilt in unserer Geometrie aber nicht der Satz von den drei Mittelloten eines Dreiecks, denn man kann Beispiele von drei Punkten A, B, C und drei Geraden a, b und c angeben, für die $AcBc = 1$; $BaCa = 1$ und $CbAb = 1$ gilt, ohne daß $abcabc = 1$ erfüllt ist[7b]).

Wir werden im folgenden den Begriff einer zu einer allgemeinen Gruppe gehörigen Geometrie noch etwas weiter fassen als bisher. *Wir wollen das etwas künstliche (und in seinen Konsequenzen nicht ohne weiteres zu über-*

[7a]) Es gilt hier eigentlich mehr als in der gewöhnlichen Geometrie, denn dort ist ja der hier in Formel (20) ausgesprochene Satz nicht immer richtig, weil zu jeder Dreiecksecke zwei Winkelhalbierende gehören, eine innere und eine äußere, und man unter den sechs Winkelhalbierenden geeignete Tripel aussuchen muß, die einem Büschel angehören. Hier sind dagegen überhaupt nur drei Winkelhalbierende vorhanden.

[7b]) Eine andere Geometrie mit endlichvielen Elementen, die Interesse verdienen dürfte, gehört zu der Gruppe, die das direkte Produkt der vollständigen Permutationsgruppe von vier Elementen mit sich selbst darstellt. Vgl. § 9.

blickende) *Kriterium mit dem Produktelement der Periode* 3 *für die Unterscheidung von Punkten und Geraden fallen lassen, und bei gegebener Gruppe nur überhaupt eine ganz beliebige Einteilung der von der Identität verschiedenen involutorischen Elemente in zwei zueinander fremde Teilmengen zugrunde legen. Wir nennen die Elemente der einen Teilmenge dann Geraden, die der andern Punkte und konstruieren uns eine zugehörige Geometrie auf Grund der Definitionen* (12) *bis* (17).

Wenn der Gedanke, zu beliebigen Gruppen auf die eben geschilderte Art Geometrien zu konstruieren, explizit auch in keiner der Arbeiten von Hjelmslev ausgesprochen ist, so bewegen sich dessen Untersuchungen doch in einer Richtung, die diese Idee implizite in sich schließt. Nach einigen Versuchen anderer Autoren, die mehr in den Anfängen stecken geblieben waren[8]), hat Hjelmslev[9]) als erster zum Beweise geometrischer Sätze in weiterem Umfang das Rechnen mit Gruppenrelationen herangezogen, und im besonderen das Rechnen mit Spiegelungsreihen, d. h. mit Gruppenrelationen, die in den involutorischen Elementen der Spiegelungen geschrieben sind. In den Arbeiten von Hjelmslev findet sich auch ein Axiomensystem, mit Hilfe dessen sich die Gruppe der kongruenten Abbildungen der gewöhnlichen euklidischen ebenen Geometrie durch innere gruppentheoretische Eigenschaften aus der Gesamtheit aller möglichen Gruppen herausheben läßt. Dies Problem der Kennzeichnung der „euklidischen Kongruenzgruppe" ist schon in älteren Arbeiten von Lie und Hilbert behandelt worden. Besonders bei Hilbert ist die Kennzeichnung durch ganz wenige einfache Axiome gegeben[9a]). Im Gegensatz zu Hjelmslev setzen Lie und Hilbert aber die Gruppe von vornherein als Gruppe von Abbildungen einer kontinuierlichen Punktmannigfaltigkeit voraus, welche auf die zweidimensionale Zahlenebene oder auf ein Teilgebiet derselben abbildbar ist, und aus ihren Axiomen ist dann zu ersehen, daß auch die Gruppe als kontinuierlich vorausgesetzt wird. Die Hjelmslevschen Untersuchungen bewegen sich mehr im Rahmen eines Problems, das wir folgendermaßen formulieren wollen: *Unter der Gesamtheit aller Gruppen sollen durch innere Eigenschaften diejenigen gekennzeichnet werden, welche äquivalent sind zu den kongruenten Abbildungsgruppen* (10), *die zu den analytischen Geometrien geometrischer Körper gehören.* Solche Gruppen können noch unendliche diskrete Gruppen oder kontinuierliche Gruppen im gewöhnlichen Sinne oder

[8]) Vgl. z. B. Herrmann Wiener, Ber. der Leipz. Ges. der Wissensch. 1890 (S. 13–23; 81–87; 245–267), sowie 1891 (S. 424–447; 644–673).

[9]) Vgl. Johannes Hjelmslev, Einleitung in die allgemeine Kongruenzlehre, Det Kgl. Danske Videnskabernes Selskab. Math.-Phys. Meddelelser, VIII. (11) (1929), sowie X. 1 (1929).

[9a]) Vgl. Hilbert a. a. O., Anhang IV.

auch noch anders geartete Gruppen sein. Alle Stetigkeitseigenschaften sind bei diesem Problem zunächst ganz ausgeschieden. In ihm haben wir, wie wir sagen können, das eigentlich elementar-geometrische Problem der Kennzeichnung der im weiteren Sinne „euklidischen" Kongruenzgruppen vor uns. Im Hinblick auf dieses Problem wird unsere allgemeine, zunächst aus einer beliebigen Gruppe heraus entwickelte Geometrie folgendermaßen weiter auszubilden sein: Wir nehmen nach und nach immer mehr Axiome zu den Gruppenaxiomen hinzu, indem wir nach und nach immer mehr Forderungen an die zugrunde gelegte Gruppe stellen. Diese Axiome werden natürlich Aussagen sein, die für die Geometrien allgemeiner geometrischer Körper richtig sind, die aber für die allgemeinsten Gruppengeometrien zunächst nicht gültig zu sein brauchen. Mit jedem solchen Zusatzaxiom werden dann in der Gruppengeometrie neue Sätze beweisbar. Die Hinzunahme neuer Axiome treiben wir so weit, bis der ganze Umfang der Elementargeometrie gewonnen ist, d. h. der ganze Umfang der zu den geometrischen Körpern gehörigen Geometrie.

Erläutern wir an einem Beispiel, wie in der Gruppengeometrie durch Annahme von Zusatzaxiomen Sätze beweisbar werden können, die vorher nicht beweisbar waren. Wir stellen z. B. an die zugrunde gelegte Gruppe und die zugrunde gelegte Einteilung ihrer involutorischen Elemente in zwei Teilmengen (Punkte und Geraden) die beiden Forderungen:

1. Für drei Gerade a, b und c gilt *nie* $abc = 1$.

2. Für drei Punkte P, Q und R gilt *immer* $PQRPQR = 1$.

Da drei Geradenspiegelungen stets eine ungleichsinnige kongruente Abbildung, also niemals die Identität, drei Punktspiegelungen aber immer als Produkt wieder eine Punktspiegelung PQR ergeben, sind beide Aussagen in der gewöhnlichen und, wie leicht zu sehen, auch in den allgemeineren Geometrien geometrischer Körper richtig. Man kann dagegen leicht Gruppen konstruieren, in denen selbst bei einer Einteilung der involutorischen Elemente nach Maßgabe des Kriteriums der Periode 3, diese Axiome oder nur eines von ihnen nicht erfüllt sind. Mit Hilfe der beiden Axiome können wir nun den folgenden Satz beweisen, der für ganz allgemeine Gruppengeometrien nicht richtig zu sein braucht: Sind drei Gerade a, b und c zu einer vierten von ihnen verschiedenen Geraden s senkrecht, so gehören sie einem Büschel an. In Formeln:

$$(21) \qquad \left.\begin{matrix} asas = 1 \\ bsbs = 1 \\ cscs = 1 \end{matrix}\right\} \rightarrow \{abcabc = 1\}.$$

Der Beweis läuft so: Nach Voraussetzung sind die Produkte $a \cdot s$, $b \cdot s$ und $c \cdot s$ involutorische Elemente. Sie können nicht gleich der Identität

n, denn aus $as = 1$ würde z. B. folgen $a = s$. Wir setzen aber $a \neq s$ raus. Ebenso kann as nicht gleich einer Geraden r sein, denn aus $as = r$ gt $asr = 1$ entgegen Axiom 1; somit sind alle drei Produkte as, bs, cs nkte. Für drei beliebige Punkte gilt aber immer Axiom 2, also $bscsasbscs = 1$. Wegen $bs = sb$, $as = sa$, $cs = sc$ können wir da- aber schreiben $assbcssabssc = 1$. Hier können wir an drei Stellen $= 1$ weglassen, und es bleibt das Gesuchte $abcabc = 1$.

Die Hjelmslevschen axiomatischen Untersuchungen gehen insofern außer- dentlich weit, als es ihr Ziel ist, die (hyperbolische) nichteuklidische ometrie und die euklidische Geometrie möglichst weitgehend gemeinsam fzubauen, ohne Axiome über das Verhalten der Parallelen in beiden Geo- etrien, ohne Anordnungs- und Stetigkeitsaxiome, unter Auflockerung der rknüpfungsaxiome und ohne Benutzung räumlicher Axiome.

Da die Behandlung der nichteuklidischen Geometrie nicht zu den Zielen r folgenden Arbeit gehört, gestalten sich manche ihrer Untersuchungen nfacher, und ihre axiomatische Analyse geht dort nicht so tief. An andern ellen werden zunächst weniger axiomatische Annahmen gemacht als bei elmslev. Die Hauptabsicht dieser Arbeit ist es, den formalen Kalkül der iegelungs- und Gruppenrechnung noch weiter auszubilden, als dies bei elmslev geschehen ist. Dieser Gesichtspunkt ist auch für die Auswahl r Zusatzaxiome maßgebend; es werden solche gewählt, die, zu den Gruppen- iomen hinzugenommen, immer noch einen übersichtlichen Kalkül für die laubten Umformungen von Gruppenrelationen gestatten. Besonders ge- gnet ist z. B. das oben erwähnte Axiom: Es gilt immer $PQRPQR = 1$. enn das heißt $PQR = RQP$ für drei beliebige Punkte. Setzt man das xiom voraus, so wird folgende Umformungsvorschrift anwendbar, die allein s den Gruppenaxiomen nicht folgt: Wenn in irgendeiner Gruppenrelation ei Punkte hintereinander vorkommen, so darf man den ersten und dritten ber den zweiten hinüber" vertauschen.

Vielleicht lassen die Ansätze dieser Arbeit es als möglich erscheinen, n Kalkül mit den Gruppenrelationen oder Spiegelungsreihen für die Be- ndlung der Elementargeometrie so auszubauen, daß er sich dem Körper- lkül der Descartesschen analytischen Geometrie als ein Konkurrenzunter- hmen an die Seite stellen kann. Im besonderen wird es darauf ankommen, ch für die räumliche Geometrie einen brauchbaren Kalkül zu entwickeln. nige erste Ansätze scheinen auch hier Erfolg zu versprechen[10]). Man

[10]) Bei der Gruppe der kongruenten Abbildungen des Raumes kann man die iegelungen an Geraden, Ebenen und Punkten auf folgende Weise durch reine ruktureigenschaften der Gruppe unterscheiden: Die Geradenspiegelungen und nur ese involutorischen Elemente (außer der Identität) lassen sich als Quadrate von uppenelementen darstellen. In der Menge M aller übrigen von der Identität ver-

(Fortsetzung der Fußnote [10]) auf nächster Seite.)

muß sich vor Augen halten, daß ein reiner Gruppenkalkül, wie er im folge den eingeführt wird, in einem wichtigen Punkte dem Körperkalkül übe legen ist. Bei der körpertheoretischen Behandlung der Geometrie verwend man wegen der Willkürlichkeit des Koordinatensystems zunächst ein Zuvi an Rechenhilfsmitteln, und dies Zuviel muß hinterher durch eine Invariante theorie, etwa mit den Methoden der Vektorrechnung wieder eliminiert werde Bei der gruppenalgebraischen Geometrie, wie sie im folgenden entwickelt wir ist von vornherein kein überzähliges Element im Kalkül enthalten, in i haben wir also die eigentlich naturgemäße Behandlung der Elementa geometrie zu erblicken. Außerdem läßt sich der Kalkül der gruppenalgebr ischen Geometrie so einrichten, daß man im wesentlichen mit der einen Grun operation der Multiplikation der Gruppenelemente auskommt. Bei der körpe theoretischen Methode haben wir zwei Grundverknüpfungen, Addition u Multiplikation (natürlich nebst deren Umkehrungen), deren Rechengeset uns nur wegen jahrelanger Gewöhnung so sehr einfach erscheinen. D Hauptaufgabe der gruppentheoretischen Methode liegt natürlich im ga Formalen, in der Auffindung einer kurzen und brauchbaren Schreibweis

Während so der Ausbau eines Kalküls das Hauptziel der Arbeit i wird das eben erwähnte Problem der Kennzeichnung der Kongruenzgrupp geometrischer Körper nur nebenbei gestreift. Im Anhang dieser Arb findet sich ein System von im wesentlichen sechs Axiomen, die die Gruppen unter allen möglichen kennzeichnen. Dies Axiomensystem unte scheidet sich dadurch von dem Hjelmslevschen, daß in ihm die nich euklidische Geometrie nicht mit in Betracht gezogen ist. Es leistet axiom tisch nicht so viel, ist aber schneller zu überschauen.

Was die Ausbildung des Gruppenkalküls angeht, so scheint dies Axiome system noch nicht geeignet, in allen seinen einzelnen Axiomen einfache u übersichtliche Umformungsvorschriften für Gruppenrelationen zu liefe *Diese Arbeit begnügt sich in ihrem Hauptteil daher damit, zu den Gruppe axiomen fünf andere Zusatzaxiome hinzuzunehmen, die einen brauchbar Kalkül gestatten.* Mit diesen Axiomen läßt sich schon ein großer Teil d Sätze der Elementargeometrie rein rechnerisch beweisen, z. B. folgen a ihnen die vier Sätze vom Schnittpunkt der drei Mittellinien, der drei Höhe der drei Mittellote und der drei Winkelhalbierenden eines Dreiecks. Au die Unabhängigkeitsbeweise für diese fünf Axiome werden mit einer Au nahme geführt.

---

schiedenen involutorischen Elemente kann man die Ebenenspiegelungen dadurch ken zeichnen, daß sie sich mit mindestens einem Element aus M zu einem Produkteleme der Periode 3 zusammensetzen lassen. Bei Punktspiegelungen geht dies nicht. V wendet man diese Kriterien, so kann man zu jeder Gruppe eine Raumgeometrie ko struieren.

Zum Schluß sei noch auf eine besondere Frage hingewiesen. Ohne weifel empfinden wir als die eigentlich inhaltvollsten und die wichtigsten nzelaussagen der Elementargeometrie die von der folgenden reinen Form: Wenn eine Reihe von geometrischen Gebilden, d. h. eine Anzahl von nkten, Geraden usw. gegeben vorliegt, und zwar derart, daß zwischen n gegebenen Punkten, Geraden usw. die und die geometrischen Lageziehungen bestehen (vereinigte Lage, Senkrechtstehen, Parallellaufen, Mittelpunkt sein" und anderes mehr), dann ist eine notwendige Folge eser Annahmen, daß auch noch diese bestimmte weitere geometrische gebeziehung gleichzeitig besteht." In Sätzen dieser Form kommt nichts n Eindeutigkeits- und Existenzaussagen vor, welche wir meist mehr als ebensache empfinden. Denn weder in den Voraussetzungen eines Satzes, r in obiger Form ausgesprochen wird, noch in den Folgerungen kommt was von eindeutiger Bestimmtheit irgendwelcher Elemente vor. Es interiert uns an solchem Satz verhältnismäßig wenig, ob etwa die vorliegende rade, die nach Voraussetzung durch zwei vorliegende Punkte geht, deren nzige Verbindungsgerade ist. Uns kommt es nur darauf an, aus gewissen geeigenschaften vorgegebener Dinge auf weitere Lageeigenschaften dieser ben Dinge zu schließen. In obigem Satz werden weiter auch keine istenzaussagen gemacht. Zunächst, was das wichtigste ist, nicht in den lgerungen. Dann aber auch nicht in den Annahmen. Wir nehmen an: enn die und die Dinge in den und den Beziehungen gegeben vorliegen ..., w. Wir machen aber keinerlei Voraussetzungen darüber, ob eine solche nfiguration in unserer Geometrie existieren kann. Der Schluß ist nur: enn sie existieren, dann .... Falls die Konfiguration nicht existiert, der tz also gegenstandslos wird, betrachten wir ihn nach der üblichen Konntion „gegenstandslos, also richtig" als richtig.

Man kann sich aus dem Gesamtaussagengebiet der Geometrie nun s Teilgebiet aller Sätze und Aussagen herausgetrennt denken, in denen hts von Eindeutigkeit und Existenz vorkommt. In dieser Teilgeometrie ne Existenz- und Eindeutigkeit sollen dann alle Sätze die soeben angebene Form haben. Z. B. wäre der Satz vom Höhenschnittpunkt des eiecks, wenn man die Existenzaussage in der Folgerung vermeiden will, zu formulieren: „Wenn vier Punkte A, B, C und H und sechs Gerade b, c, p, q und r gegeben vorliegen derart, daß A sich mit b, c und p vereinigter Lage befindet, B aber mit a, c und q, C mit a, b und r d endlich H mit p und q, und wenn dabei die Paare a, p, sowie b, q, wie c, r jedes für sich senkrecht sind, dann ist eine notwendige Folge, ß auch H und r sich in vereinigter Lage befinden." Wenn wir Anlnungs- und Stetigkeitsbeziehungen ausschließen und uns nur auf solche ziehungen beschränken, wie wir sie im folgenden durch unsere Gruppen-

relationen zwischen den Spiegelungen ausdrücken können, so haben all Sätze unserer Teilgeometrie die Form: „Gelten die und die Gruppen relationen, so soll diese weitere Relation eine notwendige Folge vo ihnen sein."

Man kann sich nun die Aufgabe stellen, diese Teilgeometrie alle elementargeometrischen Sätze ohne Eindeutigkeits- und Existenzaussage (sowie ohne Anordnungs- und Stetigkeitsbeziehungen), die wohl immer noc den wesentlichsten Bestandteil der Elementargeometrie darstellt, für sic axiomatisch zu begründen. Man kann dabei nicht vom Gruppenbegriff un den zugehörigen Geometrien ausgehen, denn in den Gruppenaxiomen stecke ja schon Eindeutigkeits- und Existenzaussagen (eindeutige Existenz eine Produktelements zu zwei gegebenen, eindeutige Existenz des inversen Ele ments usw.), die sich dann auf ähnliche Aussagen der Geometrie übe tragen. Z. B. gilt in jeder Gruppengeometrie der Satz: Es gibt zu zwe Punkten P und Q genau ein Element α, das entweder Punkt oder Gerad ist, und für das $PQ\alpha Q = 1$ ist. In der Tat ist $\alpha = QPQ$ das einzig Element, das diese Gleichung erfüllt, und zwar ist dieses Element ein vo der Identität verschiedenes involutorisches Element. Damit hängt zusamme daß in der in dieser Arbeit durch Zusatzaxiome weiterentwickelten Gruppe geometrie auch reine Existenzaussagen gefolgert werden können. So zu Beispiel im § 6 der eben vor Satz 10 genannte Satz und der Satz: Zw senkrechte Gerade haben stets genau einen Schnittpunkt.

Man kann nun alle Eindeutigkeits- und Existenzaussagen ausschalte indem man sich ganz vom Gruppenbegriff loslöst. Man definiert die B ziehungen „vereinigte Lage", „Senkrechtstehen" usw. nach (12) bis (1 durch gewisse in großen und kleinen Buchstaben (Punkten und Gerade geschriebene Reihen, die mit $= 1$ enden, wobei dieses Symbol ebenso w die Reihen überhaupt keine weitere Bedeutung haben. Man kann dann d Umformungsregeln, mittels derer man früher auf Grund der Gruppenaxiom aus solchen gegebenen Reihen neue gewinnen konnte (vgl. das Beispiel (18 axiomatisch einführen, so daß man hinterher doch genau so rechnen da wie mit Gruppen, ohne daß jetzt aber etwas über irgendwelche Existe oder Eindeutigkeit ausgemacht ist. Zu diesen gruppentheoretischen Um formungsgesetzen für die Reihen müssen dann, um unsere Teilgeometr aufzubauen, noch weitere Gesetze als Zusatzaxiome hinzugenommen werde Wir setzen dann nicht voraus, daß aus der Reihe $\alpha\beta = 1$ für zwei El mente α, β deren Identität folgt. Eine vollständige axiomatische Grun legung dieser Teilgeometrie ohne Eindeutigkeits- und Existenzaussagen s in einer späteren Arbeit gegeben werden. Das wesentliche an einer solch Geometrie ist, daß sich aus ihr die Begriffe „es gibt" und „alle" in g wissem Sinne eliminieren lassen, so daß ihre Schlüsse sich ganz im Rahm

r finiten Aussagenlogik[10a]) vollziehen. Die Axiomatik einer solchen Geo-
etrie ist durch äußerste Reinheit der Methode gekennzeichnet. Eindeutig-
it und Existenz kann man am Schluß des ganzen Aufbaus der Geometrie
neinbringen, wenn man will, indem man geeignete ideale Elemente einführt.

Die Eliminierung von Existenz- und Eindeutigkeitsaussagen spielt auch
i Hjelmslev schon eine gewisse Rolle. Z. B. setzt er ausdrücklich das
xiom: Zwei Punkte bestimmen genau eine Verbindungsgerade, das bei
ilbert allen weiteren Untersuchungen zugrunde gelegt wird, nicht voraus.
ber es kommen bei ihm in den Axiomen an andern Stellen Existenz- und
indeutigkeitsannahmen vor. Außer den Voraussetzungen, die im Gruppen-
griff enthalten sind, wird bei ihm z. B. noch angenommen, daß es durch
nen gegebenen Punkt zu einer gegebenen Geraden genau ein Lot gibt,
rner, daß es zu zwei Punkten mit Verbindungsgerade immer ein Mittellot
bt, und schließlich, daß es zu zwei, auf derselben oder auf verschiedenen
eraden gelegenen kongruenten Punktreihen, welche einen Punkt gemeinsam
aben, stets eine Gerade gibt, durch deren Spiegelung sie ineinander über-
führt werden.

Die Ausschaltung sämtlicher Eindeutigkeits- und Existenzaussagen ist
elleicht auch vom Standpunkt der aktuellen Fragen der logischen Grund-
gung der Mathematik von Interesse, etwa für das Problem, für gewisse
eilgebiete der Geometrie direkte Widerspruchslosigkeitsbeweise zu führen,
ne auf die Widerspruchslosigkeit der Arithmetik zurückzugehen[11]).

§ 1.

Spiegelungsreihen in der gewöhnlichen euklidischen Geometrie.

Man pflegt bei gewissen Untersuchungen der Elementargeometrie mit
orteil mit den Gruppenelementen der einzelnen kongruenten Abbildungen
Bewegungen und Umlegungen[12])) der Ebene zu rechnen. Ist z. B. das
esultat der Hintereinanderausführung der Abbildungen $\mathfrak{A}$ und $\mathfrak{B}$ dasselbe
ie das der Hintereinanderausführung der Abbildungen $\mathfrak{P}$ und $\mathfrak{Q}$, so schreibt
an die Gruppenrelation $\mathfrak{B}\mathfrak{A} = \mathfrak{Q}\mathfrak{P}$ oder, wenn man von rechts erst mit
em Inversen $\mathfrak{P}^{-1}$ von $\mathfrak{P}$ und dann mit dem Inversen $\mathfrak{Q}^{-1}$ von $\mathfrak{Q}$ multi-
iziert:

$$\mathfrak{B}\mathfrak{A}\mathfrak{P}^{-1}\mathfrak{Q}^{-1} = 1, \qquad (2)$$

[10a]) Vgl. Hilbert-Ackermann, Grundzüge der theoretischen Logik. (Berlin: Julius
ringer 1928). Grundlehren der math. Wissensch. XXVII, Kapitel I.

[11]) Vgl. etwa die Ausführungen von Dehn in dem Buche Pasch-Dehn, Vorlesungen
er neuere Geometrie (Berlin: Julius Springer 1926), S. 260–261.

[12]) Umlegungen sind solche kongruente Abbildungen, die den Winkelsinn um-
hren, wie z. B. die Spiegelungen an Geraden.

wo 1 die Identität bedeutet. Wir machen nun eine erste Festsetzung: *Wir wollen jede Gruppenrelation zwischen kongruenten Abbildungen immer auf diese Normalform bringen, in der alle Glieder auf die linke Seite geschrieben sind und rechts die Identität steht, und dann wollen wir in der Relation das* $=1$ *als überflüssig nicht mehr mitschreiben.* Wenn wir somit sagen: Es gilt die Reihe $\mathfrak{B}\mathfrak{A}\mathfrak{P}^{-1}\mathfrak{Q}^{-1}$, so heißt das, es gilt (22); wenn wir sagen: Es gilt die Reihe $\mathfrak{B}\mathfrak{A}\mathfrak{B}\mathfrak{P}^{-1}\mathfrak{Q}\mathfrak{P}^{-1}$, so ist damit z. B. wegen des fortgelassenen $=1$ gleichbedeutend, daß $\mathfrak{B}\mathfrak{A}\mathfrak{B} = \mathfrak{P}\mathfrak{Q}^{-1}\mathfrak{P}$ ist.

Die einfachsten Elemente der Gruppe der kongruenten Abbildungen sind von einem gewissen Standpunkt aus die *involutorischen*, die mit ihrem Inversen gleich sind, oder die, zweimal hintereinander ausgeführt, die Identität ergeben. Die von der Identität verschiedenen involutorischen kongruenten Abbildungen bezeichnen wir zunächst mit griechischen Buchstaben $\beta, \gamma, \varrho, \tau, \ldots$ usw. Für ein solches Element gelten dann die Reihen $\beta\beta$ und $\beta\beta^{-1}$, d. h. $\beta\beta = 1$ und $\beta\beta^{-1} = 1$. In Gruppenrelationen mit nur involutorischen Elementen können wir die Exponenten -1 wie in (22) entbehren, da ja die Inversen gleich den Elementen selbst sind. Die einzigen involutorischen kongruenten Abbildungen der Ebene sind die Spiegelungen an Geraden und die Spiegelungen an Punkten. Nun läßt sich jede kongruente Abbildung der Ebene erzeugen durch Hintereinanderausführung von solchen Spiegelungen. Wir können daher jede Abbildung $\mathfrak{A}$ ersetzen durch ein Produkt gewisser Spiegelungen, etwa $\mathfrak{A} = \beta\gamma\varrho\lambda \ldots \tau$.

Das Ausdrücken aller Abbildungen durch Produkte von involutorischen Abbildungen und die Zurückführung aller Gruppenrelationen auf solche mit lauter involutorischen Elementen bringt gewisse rechnerische Vorteile mit sich: Gilt z. B. eine solche *Reihe* — so wollen wir eine Gruppenrelation lauter involutorischer Elemente in der Normalform jetzt immer nennen —:

$$\beta\lambda\gamma\varrho\sigma\tau, \tag{23}$$

so kommt es erstens nicht darauf an, ob man die Reihe so oder in umgekehrter Reihenfolge $\tau\sigma\varrho\gamma\lambda\beta$ schreibt, denn aus $\beta\lambda\gamma\varrho\sigma\tau = 1$ folgt, wenn man nacheinander von rechts mit $\tau, \sigma \ldots$ usw. multipliziert: $1 = \tau\sigma\varrho\gamma\lambda\beta$. Zweitens kommt es nicht darauf an, daß man die Reihe (23) gerade mit dem Element β beginnt. Man kann mit irgendeinem Element, etwa ϱ, anfangen, schreibt dann die folgenden Buchstaben σ, τ der Reihe weiter und setzt die am Anfang weggelassenen Elemente β, λ, γ ans Ende, dann allerdings in der richtigen Reihenfolge. Man bekommt so $\varrho\sigma\tau\beta\lambda\gamma$. Daß diese Reihe mit (23) gleichbedeutend ist, erkennt man, wenn man (23) $\beta\lambda\gamma\varrho\sigma\tau = 1$ erst von rechts und links mit β, dann ebenso mit λ und dann mit γ multipliziert. Man kann eine jede Reihe also in einen

geschlossenen Zyklus geschrieben denken (vgl. Fig. 1), und es kommt dann nur auf die Reihenfolge der Elemente in diesem Zyklus unabhängig von dem Umlaufssinn an. Z. B. gestattet die Reihe $\beta\gamma\varrho\beta\gamma\varrho$ die sechs Schreibweisen:

$$24) \quad \begin{cases} \beta\gamma\varrho\beta\gamma\varrho; & \varrho\gamma\beta\varrho\gamma\beta, \\ \gamma\varrho\beta\gamma\varrho\beta; & \gamma\beta\varrho\gamma\beta\varrho, \\ \varrho\beta\gamma\varrho\beta\gamma; & \beta\varrho\gamma\beta\varrho\gamma. \end{cases}$$

Wir wollen nun die beiden Arten von involutorischen Elementen, die Spiegelungen an den Geraden und die an den Punkten, unterscheiden. Wir verwenden weiter griechische Buchstaben, wenn es gleichgültig ist, ob wir Spiegelungen an Punkten oder Geraden dafür einsetzen; wollen wir aber ausdrücklich eine der beiden Arten nehmen, so bezeichnen wir die Spiegelungen an Geraden mit kleinen lateinischen Buchstaben $a, b, c, d, g, \ldots$, die Spiegelungen an Punkten aber mit großen lateinischen Buchstaben P, Q, S, T usw. $aPaP$ heißt jetzt also: Die Aneinanderreihung der Spiegelungen an der Geraden a, am Punkt P, dann wieder an der Geraden a und nochmals am Punkte P ergibt die Identität. Man kann dafür auch $PaPa$ oder $Pa = aP$ schreiben: Die Spiegelungen a und P sind vertauschbar.

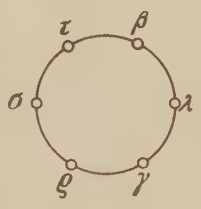


Fig. 1.

§ 2.

Darstellung geometrischer Beziehungen durch Spiegelungsreihen in der gewöhnlichen euklidischen Geometrie.

Die wichtigsten Lagebeziehungen zwischen Punkten und Geraden der euklidischen Ebene lassen sich, wie schon in der Einleitung betont wurde [(12) bis (17)], sehr einfach darstellen durch Gruppenrelationen zwischen den Spiegelungen an diesen Punkten und Geraden[13]). Wir stellen vorweg die Tatsachen fest: Schneiden sich die Geraden a und b unter dem von a nach b gerechneten Winkel φ, so ist ab die Drehung um den doppelten Winkel 2φ. Sind a und b parallel, so ist ab die Translation senkrecht zu ihnen um den doppelten Abstand in der Richtung von a nach b. Für zwei Punkte P und Q ist PQ die Translation um den doppelten Vektor, der von P nach Q führt.

Wir geben jetzt eine Liste von einfacheren Spiegelungsreihen nebst den geometrischen Lagebeziehungen, für die sie charakteristisch sind. Erläuterungen dazu schließen wir an. Überall können wir als Repräsen-

[13]) Die hier angegebenen Spiegelungsreihen finden sich bei Hjelmslev, a. a. O. Nur (11) und (17) sind dort nicht verwendet.

tanten für die Punkte und Geraden die zugehörigen Spiegelungen ein treten lassen.

1. $aPaP$ — Punkt und Gerade in *vereinigter Lage.*
2. $abab$ — *Senkrechte Gerade.*
3. $PSQS$ — S ist *Mittelpunkt* der Strecke PQ.
4. $adbd$ — Es sind entweder a und b parallel und d ist ihr *Mittellinie*, oder a und b schneiden sich und d is Winkelhalbierende. [Wir wollen auch im letzteren Falle die Winkelhalbierende als Mittellinie von und b bezeichnen.]
5. $PaQa$ — a ist *Mittellot* der Strecke PQ.
6. $aSbS$ — a und b sind notwendig parallel und S liegt au der Mittellinie.
7. $abcabc$ — *Entweder gehen a, b und c durch denselben Punk oder sind alle drei parallel.*
8. $PQRS$ — P, Q, R, S bilden ein *Parallelogramm*, bei dem und R ein Paar von Gegenecken sind. Das Parallelo gramm darf im Grenzfall in eine Gerade hineinfalle
9. $abdg$ — Entweder alle vier Gerade gehen durch einen Punk und der *Winkel $\overrightarrow{ab}$ ist gleich dem Winkel $\overrightarrow{gd}$* Oder alle vier Gerade sind parallel und der Ab stand $\overrightarrow{ab}$ ist gleich dem Abstand $\overrightarrow{gd}$.
10. abZ — a und b sind senkrecht und Z ist ihr Schnittpunkt
11. $abZabZ$ (aber nicht abZ) — *a und b sind parallel* und Z ist ein ganz beliebi zu den Geraden gelegener Punkt.
12. $aPQaPQ$ — a ist senkrecht zur Verbindungslinie von P und Q
13. $aPQaQP$ — a ist parallel zur Verbindungslinie von P und Q
14. $ababab$ — a und b bilden den Winkel $\pi/3$.
15. $\underbrace{abab\ldots ab}_{n\text{-mal}}$ (aber nicht $\underbrace{abab\ldots ab}_{p\text{-mal}}$, wo p Teiler von n) — Der von a und b gebildete spitze Winkel ist eine der möglichen Winkel $\frac{m}{n}\pi$ zwischen 0 und π 2 wobei m zu n teilerfremd ist. Wir sagen: a und bilden einen *Winkel der Periode* $2n$.
16. $\underbrace{PSPS\ldots PS}_{p\text{-mal}}\underbrace{QSQS\ldots QS}_{q\text{-mal}}$ heißt: S teilt die Strecke $\overrightarrow{PQ}$ im Ver hältnis $q:p$

7. | $abcdZabcdZ$ (aber nicht $abcdZ$) | Der Winkel $\overrightarrow{ab}$, den die Geraden a und b bilden, ist dem Winkel $\overrightarrow{dc}$, den die Geraden d und c bilden, gleichsinnig kongruent. Die sonstige Lage der beiden Geradenpaare sowie die Lage des Punktes Z zu ihnen ist ganz willkürlich.

Bevor wir auf die Einzelheiten dieser Tabelle eingehen, schicken wir och folgendes voraus: Bekanntlich kann man eine kongruente Abbildung $\mathfrak{A}$ rerseits wieder einer kongruenten Abbildung $\mathfrak{B}$ unterwerfen. $\mathfrak{A}$ bestimmt a eine Zuordnung von Punktepaaren (Original → Bild). Wenn man nun diese uordnung der Paare bei der Abbildung $\mathfrak{B}$ mitnimmt, so entsteht aus ihr ine neue Abbildung $\mathfrak{C}$, die „*mittels* $\mathfrak{B}$ *transformierte Abbildung* $\mathfrak{A}$". ekanntlich stellt sie sich durch

$$\mathfrak{C} = \mathfrak{B}^{-1}\mathfrak{A}\mathfrak{B}$$

ar. Ist $\mathfrak{B}$ als Produkt involutorischer Abbildungen dargestellt,

$$\mathfrak{B} = \beta\gamma\varrho\ldots\sigma\tau,$$

o ist

$$\mathfrak{B}^{-1} = \tau\sigma\ldots\varrho\gamma\beta$$

nd

$$\mathfrak{C} = \tau\sigma\ldots\varrho\gamma\beta\mathfrak{A}\beta\gamma\varrho\ldots\tau\sigma.$$

s ist dann $\tau\mathfrak{A}\tau$ die mit der Spiegelung τ transformierte Abbildung $\mathfrak{A}$, eiter $\sigma\tau\mathfrak{A}\tau\sigma$ die Abbildung $\mathfrak{A}$, die mit der aus σ und τ entstehenden ewegung $\sigma\tau$ transformiert ist, usw. Mittels dieser Einsichten über die ransformation von Abbildungen können wir leicht die Richtigkeit der in nserer Tabelle enthaltenen Behauptungen erkennen: Wir beginnen mit 3., a 1. und 2. uns schon bekannt sind. 3. schreiben wir als $P = SQS$. ie Spiegelung am Punkte Q transformiert mit der Spiegelung am Punkte S rgibt die Spiegelung an dem Punkt, der aus Q durch Spiegelung an S ntsteht. Dieser Punkt soll nach 3. aber P sein. Also ist tatsächlich S der ittelpunkt der Strecke PQ, wie behauptet. Analog behandelt man 4., 5., 6. ie Formel 7. schreibt man als $abca = cb$. Bilden hier b und c den inkel φ, so soll die zugehörige Drehung $\overrightarrow{bc}$ um den doppelten Winkel 2φ ittels der Spiegelung a transformiert die Drehung cb um den entgegen-esetzten Winkel -2φ ergeben. Der Schnittpunkt von b und c als das emeinsame Zentrum beider Drehungen muß dann bei der Spiegelung a st bleiben, d. h. a geht durch den Schnittpunkt von b und c hindurch. nd b und c in 7. aber parallel, so soll nach $abca = cb$ durch die piegelung a die Translation bc übergehen in die entgegengesetzte Trans-tion. a muß dann senkrecht zur Translationsrichtung, also parallel zu und c sein. Im Fall 7 bilden ab und c also immer ein *Büschel im*

projektiven Sinne. Wir wollen die drei Geraden, da ihr Produkt abc ur nach § 1 (24) auch jedes der weiteren Produkte bca, cba usw. ein i volutorisches Element darstellt, dann auch einfach als ein *involutorisch Geradentripel* bezeichnen.

Die Richtigkeit von 8. erkennen wir leicht aus $PQ = SR$, wenn w bedenken, daß die Hintereinanderreihung PQ zweier Punktspiegelunge immer eine Translation um die doppelte Strecke $\overrightarrow{PQ}$ ergibt. 9. folgt leic aus $ab = gd$. 10. können wir in den Formen $a = bZ$, $b = aZ$ ur $Z = ab$ schreiben. Da aa, bb und ZZ gilt, folgen die Reihen $bZbZ$ $aZaZ$ und $abab$, woraus nach 1. und 2. die Behauptung folgt. 11. od $ZabZ = ba$ ist sehr wichtig. Da Z eine eigentliche Bewegung ist, d den Umlaufssinn der Winkel enthält, kann durch Transformation mit nur dann aus einer Drehung ab die entgegengesetzte ba entstehen, wen eine Drehung um π vorliegt, d. h. wenn a und b senkrecht sind. Dann i ab mit ba identisch, und es muß weiter Z der Schnittpunkt von a und sein. Wir haben dann den Fall 10., der bei 11. aber ausdrücklich au geschlossen wird. Also bleibt nur die Möglichkeit, daß a und b parall sind. Dann ist 11. aber für beliebige Lage von Z richtig. Denn durc Spiegelung an einem Punkte entsteht aus einer Translation ab immer d entgegengesetzte ba. *Es ist also besonders zu betonen, daß wir de Parallelismus zweier Geraden durch* 11. *mit einem beliebigen Hilfspun zum Ausdruck bringen können. Gilt* 11. *für einen Punkt Z, so gilt no wendig $abTabT$ für jeden Punkt T, und es gilt dabei nie abT.*

Schreiben wir 12. und 13. in die Formen $aPQa = QP$ und $aPQa = PQ$ und bedenken wir, daß PQ Translation ist, so können wir die Bedeutun dieser Reihen ähnlich wie die für 7. erkennen. In 14. und 15. können a und nicht parallel sein, denn eine Translation ab, beliebig vervielfacht, gi nie die Identität. a und b bilden also in 15. einen Winkel φ, für de $2n\varphi$ gleich einem Vielfachen von 2π ist, während $2p\varphi$ nicht gleich eine solchen Vielfachen ist. Daraus folgt die Behauptung. In 16. muß d q-mal vervielfachte Translation QS gleich der p-mal vervielfachten Tran lation SP sein, woraus die Behauptung folgt. In 17. muß $abcdZ$ i volutorisch sein, und zwar eine Bewegung, da die Geradenspiegelungen der geraden Anzahl 4 vorkommen. Wegen des Zusatzes ist $abcdZ$ som gleich einer gewissen Punktspiegelung T. Es gilt dann $abcd = TZ$, ist also $abcd$ eine Translation. Wenn ab und cd beide Drehungen sin (im allgemeinen um verschiedene Punkte), so müssen die Drehwinkel en gegengesetzt gleich sein, damit durch Hintereinanderreihung eine Tran lation entsteht. Also sind die Winkel $\overrightarrow{ab}$ und $\overrightarrow{dc}$ gleichsinnig kongruen Wenn ab eine Translation ist, so muß auch cd Translation sein, dan

las Produkt gleichfalls eine Translation wird. Dieser Fall ist aber ein pezialfall der Winkelkongruenz, indem beide Winkel 0 sind. *Hier ist ieder wesentlich, daß sich die Kongruenz von Winkeln mittels eines ganz eliebigen Hilfspunktes Z zum Ausdruck bringen läßt.*

Neben den einfachen Reihen der Tabelle seien noch einige kompliziertere ebst ihrer geometrischen Bedeutung angegeben, wobei wir auf eine genauere Diskussion verzichten:

1. $A_1 M A_2 M \dots A_r M$ heißt: M ist Schwerpunkt des mit lauter gleichen lassen versehenen Punktsystems $A_1, A_2, \dots, A_r$.

2. $a h a b c h b c$ bedeutet für $(abc)^2 \neq 1$, daß h die auf die Seite a efällte Höhe des Dreiseits a, b, c ist.

3. $a b c f c b a f$ bedeutet für $(abc)^2 \neq 1$, daß f Verbindungslinie der ußpunkte der beiden Höhen ist, die im Dreiseit a, b, c zu den Seiten a nd c gehören, für $(abc)^2 = 1$ aber, daß a, b und c Winkelhalbierende in inem Dreieck mit der b gegenüberliegenden Seite f sind.

4. $a b c f g a f g b c$ bedeutet im allgemeinen Fall (es bestehen mehrere usnahmen), daß a die Verbindungslinie der Schnittpunkte der beiden eradenpaare b, c und f, g ist.

Neben den in der Tabelle angeführten Reihen gibt es nun noch eine esonders wichtige Art von Reihen, nämlich die *identischen Reihen.* Wähend die Reihen der Tabelle immer nur für Punkte und Geraden in beonders ausgezeichneten Lagebeziehungen erfüllt sind, gibt es auch Reihen, ie bei beliebiger Lage der Punkte und Geraden erfüllt sind. Das einfachste eispiel, abgesehen von trivialen, wie $b b$ und $P a a Q Q P$, ist die Reihe

25) $$P Q R P Q R.$$

n der Tat ist die mit P transformierte Translation $Q R$ immer gleich der ntgegengesetzten $R Q$, d. h. $P Q R P = R Q$. Ein weiteres Beispiel ist

26) $$P Q R S Q P S R.$$

iese Reihe besagt wegen $P Q R S = R S P Q$, daß zwei beliebige Transtionen $P Q$ und $R S$ vertauschbar sind. Aber auch identische Reihen, die erade enthalten, gibt es, z. B.

27) $$a b c a b c R a b c a b c R.$$

s ist nämlich $a b c$ wie jede Bewegung, die durch eine ungerade Anzahl on Geradenspiegelungen entsteht, eine Umlegung. Nun ist bekannt, daß ch jede Umlegung zusammensetzen läßt aus einer Spiegelung an einer eraden p und einer Translation in Richtung von p (eventuell um die trecke 0). [Jede Umlegung besitzt eine solche Fixgerade p.] Daraus sieht man, daß das Quadrat einer Umlegung und hier speziell $a b c a b c$

immer einer Translation äquivalent ist (die sich eventuell auf die Identität reduzieren kann). Diese Translation läßt sich sicher auch durch zwei geeignete Punktspiegelungen ST (mit eventuell zusammenfallenden Punkten S und T) erzeugen. Wir haben dann statt (27) $STRSTR$, was wegen (25) wirklich immer erfüllt ist. Nimmt man die zwei Translationen $abcab$ und $bcabca$, so kann man mittels des Theorems (26) über die Vertauschbarkeit der Translationen eine identische Reihe herstellen, die nur drei Gerade enthält, nämlich

(28) $$abcabcbcabcacbacbaacbacb.$$

Läßt man die zwei ziemlich am Schluß stehenden nebeneinander auftretenden a weg, so erhält man eine Reihe mit 22 Gliedern:

(29) $$abcabcbcabcacbacbcbacb.$$ [14])

§ 3.

Weitere Festsetzungen über die Schreibweise.

Die meisten Behauptungen der Tabelle des vorigen Paragraphen bleiben sinnvoll, wenn von den in einer Reihe vorkommenden Geraden einige zusammenfallen oder wenn ebenso die Punkte unter sich zusammenfallen. Z. B. erhalten wir aus 8. für $Q = S$ die Beziehung 3. Das Parallelogramm fällt dann in die Strecke PR, zwei Ecken kommen in den Mittelpunkt S von P und R zu liegen. Ebenso erhält man für $P = Q$ aus 8. $R = S$ und man bekommt ein Parallelogramm zweier aufeinanderfallender Strecken und zweier Seiten der Länge Null. Aus 3. erhalten wir für $P = Q$ die Relation $PS = SP$, die nur für $P = S$ möglich ist. Alle Punkte fallen dann zusammen. Aus 5. erhalten wir für $Q = P$ als Grenzfall des Mittellotes die Bedingung der vereinigten Lage. Ebenso bleiben die angegebenen identischen Reihen (6) bis (10) bei zusammenfallenden Punkten und Geraden richtig. Gleiches gilt von der Reihe 17, die z. B. für $c = d$, d. h. für den Fall, daß der Winkel $\overrightarrow{cd}$ gleich 0 wird, für die Geraden a und b der kongruenten Winkels die Aussage 11. des Parallelismus als Spezialfall liefert. Eine Ausnahme macht nur Nr. 2 der Tabelle. In 2. ist $abab = 1$ für zusammenfallende Gerade a und b noch richtig. Zusammenfallende Gerade wird man aber schwerlich als Grenzfall senkrechter auffassen.

---

[14]) Es dürfte die Frage von Interesse sein, ob für drei Gerade a, b, c die Reihe (29) außer den trivialen Reihen wie $aabbcc$ oder $abcaacba$ und den aus (29) durch triviale Umformungen abgeleiteten Reihen die einzige identische Spiegelungsreihe ist. Überhaupt dürfte es wünschenswert sein, eine Übersicht über alle identischen Spiegelungsreihen überhaupt zu bekommen.

Wir setzen jetzt fest: *Wenn wir im folgenden Reihen als gültig hinschreiben, so sollen in ihnen, wenn nicht anders vermerkt, die mit verschiedenen Buchstaben vorkommenden Geraden sowohl verschieden als gleich sein dürfen, und Entsprechendes soll für die Punkte unter sich gelten.* Bei 2. werden wir also jetzt ausdrücklich den Vermerk $\{a \neq b\}$ oder $ab \neq 1$ beifügen. Was wir immer voraussetzen, ist natürlich, daß die beiden Arten involutorischer Elemente nicht mit dem involutorischen Element der Identität zusammenfallen.

Nun noch einige weitere Festsetzungen über die Schreibweise. Erstens werden wir, wenn in einer Reihe gewisse Teilfolgen von Elementen mehrfach hintereinander vorkommen, zur Abkürzung die Schreibweise mit Exponenten benutzen. Z. B. schreiben wir für $aPaP$ auch $(aP)^2$, für $abcabc$ auch $(abc)^2$, für eine n-malige Folge $abab \ldots ab$ von ab auch $(ab)^n$, für $abcabcPQPQPQabcabcPQPQPQ$ auch $((abc)^2(PQ)^3)^2$. Unsere identische Reihe (27) können wir jetzt in die Form $((abc)^2R)^2$ setzen. Weitens ist es bei einer gegebenen Reihe oft von Bedeutung, daß gewisse Teilfolgen der Reihe für sich zusammen einem involutorischen Element äquivalent sind. Sei z. B. gegeben die Reihe $PadfgQgfrSSqbfbsa$. Dann ist SS äquivalent dem involutorischen Element der Identität. Die Identität SS können wir aber aus der Reihe einfach weglassen. Weiter ist bfb äquivalent einer Geraden, denn es ist die mit b transformierte Spiegelung f. Ebenso ist $fgQgf$ die mit fg transformierte Spiegelung an Q, also einer Punktspiegelung äquivalent. Es ist für das Überschauen von Reihen bequem, sich die Zusammenfassung von Teilfolgen zu neuen involutorischen Elementen durch Klammern anzumerken, und zwar schließen wir eine Teilfolge, die einem Punkt äquivalent ist, in eckige Klammern ein: $[fgQgf]$, eine Teilfolge, die einer Geraden äquivalent ist, aber in Winkelklammern $\langle\ \rangle$, also $\langle bfb\rangle$. Da es in Reihen nicht darauf ankommt, mit welchem Glied man sie beginnt, kann man mit der Klammerzusammenfassung auch vom Ende der Reihe zum Anfang überspringen. Z. B. schreiben wir, da aPa ein Punkt ist, in unserer Reihe $Pa]fg \ldots bs[a$. Insgesamt schreiben wir unsere Reihe also etwa in der Form

$$Pa]d[fgQgf]rq\langle bfb\rangle s[a.$$

Diese Klammerschreibweise dient nur zur Erleichterung der Übersicht und braucht keineswegs immer ausgeführt zu werden. Sie ist ja auch nicht immer eindeutig. Z. B. kann man für $aPaPaPb$ schreiben $\langle aPaPa\rangle Pb$ oder auch $a[PaPaP]b$ oder $[aPa]\langle PaP\rangle b$.

Die runden Klammern ( ) reservieren wir für die oben angedeutete abkürzende Schreibweise mit Potenzen. Sie können natürlich neben den anderen Klammern vorkommen. Z. B. läßt sich $tababPbababasf$ in der Form $t[(ab)^3P(ba)^3]sf$ schreiben.

Die Zusammenfaßbarkeit von Teilfolgen zu Punkten und Geraden kann man nicht nur erschließen im Falle der Transformation einer Spiegelung mittels einer anderen kongruenten Abbildung, sondern auch aus anderweitig als gültig angenommenen Reihen. Gelten z. B. gleichzeitig die Reihen

$$(abc)^2 \quad \text{und} \quad Rabcfgh$$

so ist abc involutorisch, kann also als eine durch ungerade Anzahl von Spiegelungen an Geraden erzeugte Abbildung nur selbst eine Spiegelung an einer Geraden sein. Wir können dann also die zweite Reihe schreiben als $R\langle abc\rangle fgh$.

Im Gegensatz zu den runden Klammern können die eckigen Klammern und Winkelklammern nach den bisherigen Festsetzungen immer auch wieder weggelassen werden. Das gilt nicht mehr, wenn wir sie über das Bisherige hinaus auch im folgenden Sinne verwenden: Wir schreiben die Tatsache, daß ein bestimmtes Produkt $\beta\lambda\gamma\varrho\sigma\tau$ einer Punktspiegelung resp. Geradenspiegelung gleich ist, auch für sich als $[\beta\lambda\gamma\varrho\sigma\tau]$ und $\langle\beta\lambda\gamma\varrho\sigma\tau\rangle$. Wir verwenden die eckigen und Winkelklammern also nicht mehr nur in Reihen, wo wir sie ebensogut weglassen können, sondern die Behauptung $[\beta\lambda\gamma\varrho\sigma\tau]$ steht z. B. für $\beta\lambda\gamma\varrho\sigma\tau\beta\lambda\gamma\varrho\sigma\tau = 1$ mit den Nebenaussagen, daß erstens nicht $\beta\lambda\gamma\varrho\sigma\tau$ gilt, daß das Produkt also nicht die Identität ergibt, und zweitens, daß $\beta\lambda\gamma\varrho\sigma\tau$ keine Geradenspiegelung ist.

Unter Verwendung dieser Festsetzung können wir die Bedingung Nr. 11 der Tabelle des § 2 für den Parallelismus auch als $[abZ]$ schreiben. Denn abZ ist als gleichsinnige kongruente Abbildung sicher keine Geradenspiegelung und wegen des ausdrücklichen Zusatzes $abZ \neq 1$ nicht die Identität. Ebenso kann man für Nr. 17 schreiben $[abcdZ]$.

Wir merken uns also: *In einer als richtig hingeschriebenen Reihe mit verschiedenen runden, eckigen und Winkelklammern dürfen runde Klammern (mit den zugehörigen Exponenten) überhaupt nie weggelassen werden, eckige und Winkelklammern dürfen im allgemeinen weggelassen werden, nur dann nicht, wenn sie die ganze Reihe umfassen. „Im Innern" einer Reihe vorkommende eckige und Winkelklammern darf man stets weglassen.*

Gilt eine Reihe $(\alpha\beta \ldots \lambda\mu \ldots \varrho\tau)^2$, so kann man, wie leicht einzusehen, in der Klammer alle Umformungen vornehmen, die man auch mit der Reihe $\alpha\beta \ldots \lambda\mu \ldots \varrho\tau$ selbst vornehmen könnte, wenn sie gültig wäre. Man kann in der Klammer die Reihenfolge aller Elemente umstellen, also $(\tau\varrho \ldots \mu\lambda \ldots \beta\alpha)^2$ schreiben oder man kann den Zyklus der Elemente in der Klammer mit einem in der Mitte stehenden Element, etwa μ, beginnen und schreiben $(\mu \ldots \varrho\tau\alpha\beta \ldots \lambda)^2$. Immer erhalten wir richtige Reihen. Gleiches gilt für die Schreibweisen $[\alpha\beta \ldots \lambda\mu \ldots \varrho\tau]$ und $\langle\alpha\beta \ldots \lambda\mu \ldots \varrho\tau\rangle$ für involutorische Elemente, die als Punkt- resp. Geradenspiegelung erkannt

ind. Man kann die letzte Reihe auch z. B. mit $\sigma\sigma$ zu $\langle\alpha\beta \ldots \lambda\sigma\sigma\mu \ldots \varrho\tau\rangle$ erweitern. Weiter kann man in $[\sigma\alpha\beta \ldots \lambda\mu \ldots \varrho\tau\sigma]$ vorn und hinten das σ gegeneinander kürzen. All das folgt durch Rückgang auf die Grunddefinition dieser Reihen. Aus $[abP]$ folgt auch $[bPa]$, $Pab]$, $[Pba]$, $[baP]$ und $[aPb]$, doch ind hier im allgemeinen die drei Punkte $abP] = [Pba]$ und $[bPa] = [aPb]$ und $Pab] = [baP]$ verschieden.

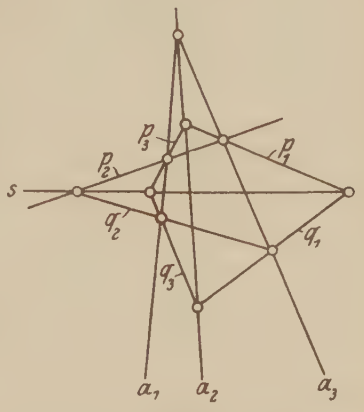


Fig. 2.

Gelten mehrere Reihen nebeneinander, o fügen wir sie wohl durch das Zeichen + nintereinander. So schreiben wir z. B. den Satz (20) der Einleitung jetzt in der Form

$$arbr + bpcp + cqaq \to (pqr)^2.$$

Das Zeichen + hat nur die Bedeutung des ogischen „und", und es kommt ihm nicht die Rolle einer neuen Operation zu. Der allgemeine Satz von Desargues (vgl. Fig. 2) lautet z. B. jetzt:

$$\left.\begin{array}{l} (s\,p_1\,q_1)^2 + (s\,p_2\,q_2)^2 + (s\,p_3\,q_3)^2 \\ \quad + (a_1\,p_2\,p_3)^2 + (a_1\,q_2\,q_3)^2 \\ \quad + (a_2\,p_3\,p_1)^2 + (a_2\,q_3\,q_1)^2 \\ \quad + (a_3\,p_1\,p_2)^2 + (a_3\,q_1\,q_2)^2 \end{array}\right\} \to (a_1\,a_2\,a_3)^2. \tag{30}$$

Statt der runden Klammern könnten wir hier überall unter Weglassung der Exponenten 2 auch Winkelklammern setzen.

§ 4.

Allgemeine Gruppengeometrien.

Wir gehen nun in dem in der Einleitung erläuterten Sinne von der gewöhnlichen euklidischen Geometrie über zu einer allgemeinen Gruppengeometrie:

Wir nehmen eine beliebige (möglicherweise auch endliche) Gruppe und teilen ihre von der Identität verschiedenen involutorischen Elemente in zwei fremde Teilmengen. Die Elemente der ersten Art nennen wir Geraden, die der zweiten Art nennen wir Punkte. Ein allgemeines Gruppenelement, das sich durch involutorische Elemente erzeugen läßt, nennen wir „konruente Abbildung". Jetzt definieren wir in Übereinstimmung mit der Tabelle des § 2, indem wir die abkürzende Schreibweise des § 3 verwenden:

1. Zwei Gerade a und b, für die nicht ab gilt, heißen *senkrecht*, wenn die Gruppenrelation $(ab)^2$ gültig ist.

2. Ein Punkt P und eine Gerade a heißen *in vereinigter Lage*, wenn $(Pa)^2$ gilt.

3. S heißt *Mittelpunkt* von P und Q, wenn $PSQS$ gilt.

4. d heißt *Mittellinie* von a und b, wenn $adbd$ gilt.

5. a heißt *Mittellot* von P und Q, wenn $PaQa$ gilt.

6. Vier Punkte P, Q, R und S bilden in dieser Reihenfolge ein *Parallelogramm*, wenn $PQRS$ gilt.

7. Drei Gerade a, b und c heißen *involutorisch oder „in einem Büschel gelegen"*, wenn $(abc)^2$ gilt.

8. S heißt *Teilpunkt* des Punktepaares P und Q (in dieser Reihenfolge) *im Verhältnis* $q:p$, wenn $(PS)^p(QS)^q$ gilt.

9. Wir sagen: a und b *„schneiden sich unter der Periode* $2n$", wenn $(ab)^n$ gilt, aber nicht $(ab)^{p'}$, wo p ein Teiler von n ist.

Nachdem wir uns so eine zu unserer Gruppe gehörige Geometrie konstruiert haben, können wir uns fragen, welche von den die obigen neun geometrischen Beziehungen enthaltenden Sätzen der gewöhnlichen Geometrie in ihr noch richtig sind. Ist z. B. der Höhenschnittpunktsatz oder der Satz von Desargues § 3 (30) richtig? Wenn nicht, welche einfachen Zusatzannahmen kann man als Axiome zu den Gruppenaxiomen eventuell hinzunehmen, damit die Sätze beweisbar werden?

Wir betonen noch, daß wir den Begriff des Parallelismus mittels Nr. 11 der Tabelle des § 2 in den allgemeinen Gruppengeometrien vorläufig nicht definieren wollen. Denn wir können in den allgemeinen Geometrien nicht zeigen: Bei gegebenen Geraden a und b zieht die Gültigkeit von Nr. 11 für einen Punkt Z die Gültigkeit für jeden Punkt nach sich, den man statt Z einsetzen kann. Wir könnten also nur einen Begriff des Parallelismus zweier Geraden relativ zu einem festen Punkt definieren oder wir könnten Definitionen einführen, wie: Zwei Gerade a, b heißen parallel, wenn es mindestens einen Punkt Z gibt, für den Nr. 11 gilt. Wir wollen aber Definitionen und Axiome möglichst vermeiden, in denen „es gibt" in dieser Weise vorkommt. Ähnliches wie vom Begriff des Parallelismus gilt auch von dem durch Nr. 17 definierten Begriff der Kongruenz von Winkeln. Wir werden später aber zeigen, daß sich Parallelismus und Winkelkongruenz schon nach Einführung ganz einfacher Zusatzaxiome im Sinne des ausgesprochenen Satzes erklären lassen. Solange wir noch nicht alle Begriffe der gewöhnlichen Geometrie definieren können, haben wir natürlich nur einen Ausschnitt aus der Geometrie.

Wir geben jetzt einige einfache Sätze an, die sich schon in den allgemeinen Gruppengeometrien, also allein auf Grund der Gruppenaxiome beweisen lassen.

Satz 1. $PQRS + RSTV \rightarrow PQVT$.

In Worten: *Bilden die vier Punkte P, Q, R und S (in dieser Reihenfolge) ein Parallelogramm und tun dies ebenfalls die vier Punkte R, S, T und V, so bilden notwendig auch die Punkte P, Q, V und T ein Parallelogramm.*

Diesen Satz pflegt man in der gewöhnlichen Geometrie auch als den *kleinen Satz von Desargues* zu bezeichnen (Fig. 3).

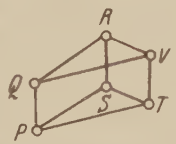


Fig. 3.

Beweis. Aus $PQRS = 1$ und $RSTV = 1$ folgt einerseits $RS = QP$ und andererseits $RS = VT$, woraus sich $QP = VT$ oder die Behauptung des Satzes ergibt.

In Übereinstimmung mit den Festsetzungen des § 3 fassen wir unsere Sätze immer so, daß sie richtig bleiben, wenn in ihnen irgendwelche Punkte unter sich oder irgendwelche Geraden unter sich zusammenfallen. Setzen wir in Satz 1 z. B. $P = R$, so haben wir nach der Tabelle dieses Paragraphen den Satz: Ist P Mittelpunkt des Paares Q, S und liegen P, S, T, V im Parallelogramm, so tun es auch die Punkte P, Q, V, T. Setzen wir außer $P = R$ noch $Q = V$, so haben wir: Ist P Mittelpunkt von QS und S Mittelpunkt von PT, so liegen $PQST$ im Parallelogramm (dieses Parallelogramm fällt in der gewöhnlichen Geometrie ganz in eine Gerade, wobei dann die Strecken PQ und ST kongruent sind).

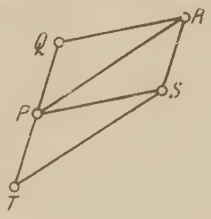


Fig. 4.

Ein weiterer Spezialfall von Satz 1 ergibt sich für $P = V$: Liegen $PQRS$ sowie $RSTP$ im Parallelogramm, so ist P Mittelpunkt von QT (vgl. Fig. 4).

Satz 2. $(abc)^2 + (abd)^2 + (acd)^2 \rightarrow (bcd)^2$.

In Worten: *Sind die Geradentripel abc, abd und acd jedes für sich involutorisch (in einem Büschel gelegen), so ist auch das Tripel bcd involutorisch.* Hier ist noch zu betonen, daß für ein Geradentripel die Aussage „involutorisch" oder „in einem Büschel gelegen" noch gar nichts mit der Aussage „alle drei gehen durch einen Punkt" oder „keine zwei haben einen Schnittpunkt" zu tun hat. (Wie bereits in der Einleitung bemerkt, ist der Satz $(abc)^2 + (abd)^2 \rightarrow (bcd)^2$ in einer allgemeinen Gruppengeometrie noch nicht beweisbar.)

Beweis. Aus $(abc)^2$ folgt $ab = cbac$. Setzen wir diesen Ausdruck für ab in die zweite vorausgesetzte Relation $(abd)^2 = abdabd = 1$ am Anfang ein, so erhalten wir $cbacdabd$. Nun ist nach der dritten Voraussetzung aber $cda = adc$. Daher können wir für $cbacdabd$ auch $caadcbd$ oder $cbdcbd$ schreiben. Das ergibt nach (24) aber das gewünschte $(bcd)^2$.

Satz 3. $acbc + (acd)^2 + (cd)^2 \to adbd$.

Ist c Mittellinie von ab, so ist eine mit a und c involutorische Gerade, die zu c senkrecht ist, gleichfalls Mittellinie von a und b.

Beweis. Setzen wir $a = cbc$ in $acdacd = 1$ an erster Stelle ein, so bekommen wir $cbdacd$ oder $bda = cdc$ oder wegen $d = cdc$ auch $bda = d$, woraus sich die Behauptung ergibt.

Satz 4. $(Pa)^2 + (RQa)^2 \to QRPaPQRa$.

Dieser Satz hat vorläufig nur rein formale Bedeutung, da z. B. die Relation $(RQa)^2$ noch nicht geometrisch erklärt ist. Wir bringen ihn nur, da wir ihn später benutzen wollen.

Beweis. Setzt man $a = PaP$ für das letzte a in $RQaRQa$ ein, so ergibt sich $RQaRQPaP$ oder $aRQPaPRQ$ oder durch Umdrehen $QRPaPQRa$, was zu beweisen war.

§ 5.

Das erste Zusatzaxiom.

Wir stellen jetzt an die unserer Geometrie zugrunde gelegte Gruppe die Forderung (Zusatzaxiom):

Axiom I. *Es gilt für drei beliebige Punkte immer $PQRPQR$, aber nie PQR.*

In der gewöhnlichen Geometrie ist dies Axiom offenbar erfüllt, denn $PQRPQR$ ist die in (25) schon erwähnte identische Reihe, und PQR gilt nie, weil eine Translation PQ nie gleich einer Punktspiegelung R sein kann.

Wir beweisen jetzt einige Sätze, die aus Axiom I folgen:

Satz 5. $PQRS + PTRV \to QTSV$.

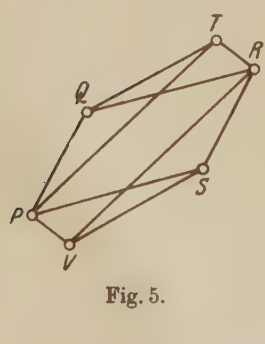


Fig. 5.

In Worten: *Liegen die Punkte $PQRS$ im Parallelogramm und tun dasselbe die Punkte $PTRV$, so liegen auch die Punkte $QTSV$ im Parallelogramm* (Fig. 5).

Beweis. Setzen wir $P = QRS$ in die zweite Reihe $PTRV$ ein, so ergibt sich $QRSTRV$. Nun ist nach Axiom I immer $RST = TSR$, man kann also in $QRSTRV$ die Buchstaben R und T „über S hinüber" vertauschen und erhält $QTSRRV$ oder das gesuchte $QTSV$.

In der gewöhnlichen Geometrie entspricht diesem Satz für den Spezialfall, daß P, Q und T auf einer Geraden liegen, der sogenannte *kleine Satz*

von *Pascal* (vgl. Fig. 6). Es ist bemerkenswert, daß er im Gegensatz zu dem im vorigen Paragraphen bewiesenen kleinen Satz von Desargues nicht aus den Gruppenaxiomen allein, sondern nur mit dem Axiom I beweisbar ist, was sich leicht durch eine Pseudogeometrie bestätigen läßt. Lassen wir im Satz 5 die Punkte T und V zusammenfallen,

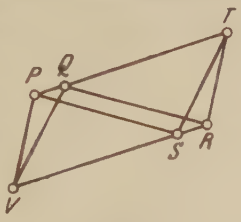


Fig. 6.

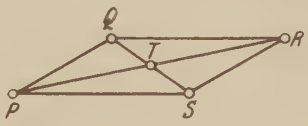


Fig. 7.

so haben wir den *Satz, daß sich die Diagonalen eines Parallelogramms halbieren* (Fig. 7).

Satz 6. $PBQB + RAQA + PMAMAM \rightarrow RMBMBM.$

Dies ist in der gewöhnlichen Geometrie der *Satz vom Schnittpunkt der drei Mittelpunktstransversalen im Dreieck* (Fig. 8): Sind drei Punkte P, Q und R gegeben und ist A Mittelpunkt von Q und R, B Mittelpunkt von P und Q und M Teilpunkt von $\overrightarrow{PA}$ im Verhältnis 2:1, so ist M notwendig auch Teilpunkt im Verhältnis 2:1 für $\overrightarrow{RB}$. Nach Satz 6 kann man hier natürlich weiter beweisen, daß M Teilpunkt im Verhältnis 2:1 auch für das Punktepaar ist, das aus dem „dritten Eckpunkt“ Q und einem Mittelpunkt von PR besteht. Für unsere allgemeine Geometrie haben wir natürlich erst das Skelett des Satzes über den Schnittpunkt der Mittelpunktstransversalen gewonnen, denn es ist bisher weder ausgemacht, daß eine Verbindungslinie zweier Dreiecksecken, etwa P und Q durch einen Mittelpunkt B des Paares PQ hindurchgeht, noch wissen wir, ob eine Verbindungsgerade von P und A durch den Teilungspunkt M hindurchgeht, das alles ganz abgesehen noch davon, daß es gleichzeitig viele Verbindungsgerade, viele Mittelpunkte und viele Teilungspunkte geben kann.

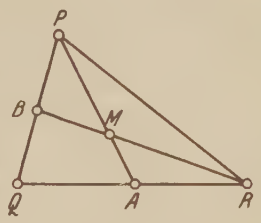


Fig. 8.

Beweis von Satz 6. Aus der zweiten Voraussetzung erhalten wir $Q = ARA$. Setzen wir das in die erste Voraussetzung ein, so erhalten wir $PBARAB$. Danach ist $P = BARAB$. Das setzen wir in die dritte Voraussetzung ein und erhalten $BARABMAMAM$. Ersetzen wir hier die unterstrichenen Tripel AMB und ABM nach Axiom I durch BMA und MBA, so erhalten wir $AARMBAAMBM$ oder $RMBMBM$, was zu beweisen war.

Wichtige Spezialfälle dieses Satzes ergeben sich für $P = B$, für $Q = M$ und für $P = R$. Wir wollen sie hier aber nicht alle ausdrücklich formulieren.

Satz 7. $\alpha P \beta P + \beta Q \gamma Q + \gamma R \delta R \dashv \delta P \varepsilon P + \varepsilon Q \varrho Q \rightarrow \alpha R \varrho R$.

Wenn wir griechische Buchstaben schreiben, so sollen nach § 1 die angeschriebenen Reihen gültig sein, einerlei ob man Punkte oder Gerade für sie einsetzt. Setzen wir für $\alpha, \beta, \gamma, \delta, \varepsilon, \varrho$ z. B. lauter Punkte ein, so haben wir den Satz (vgl. Fig. 9). *Ist P Mittelpunkt sowohl von $\{\alpha, \beta\}$ wie von $\{\delta, \varepsilon\}$ und Q Mittelpunkt sowohl von $\{\beta, \gamma\}$ wie von $\{\varepsilon, \varrho\}$, ist ferner R Mittelpunkt von $\{\gamma, \delta\}$, so ist notwendig R auch Mittelpunkt von α und ϱ.*

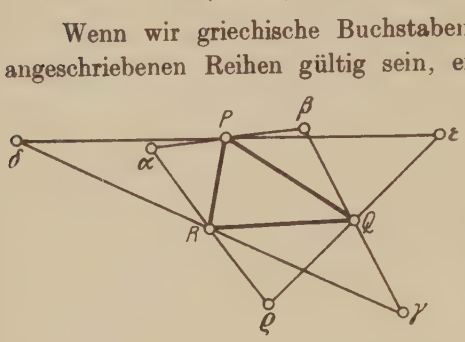


Fig. 9.

Ähnlich ist der Satz auszusprechen, wenn alle $\alpha, \beta, \gamma, \delta, \varepsilon, \varrho$ oder einige von ihnen Gerade sind.

Beweis. Aus der ersten, dritten und fünften Voraussetzung folgt $\beta = P\alpha P$, sowie $\gamma = R\delta R$ und $\varepsilon = Q\varrho Q$. Setzen wir das in die zweite und vierte Voraussetzung ein, so erhalten wir $P\alpha PQR\delta RQ$ und $\delta PQ\varrho QP$. Aus der letzten dieser beiden Reihen gewinnen wir $\delta = PQ\varrho QP$. Setzen wir das in die erste ein, so haben wir $P\alpha PQRPQ\varrho QPRQ$. Hier können wir die beiden unterstrichenen Punkttripel nach Axiom I ersetzen durch PRQ und QRP, so daß wir erhalten $P\alpha PPRQQ\varrho QQRP$ oder $\alpha R\varrho R$, woraus die Behauptung folgt.

Satz 8. Es gilt für vier beliebige Punkte immer $PQRSQPSR$.

In der gewöhnlichen Geometrie ist dieser Satz das *Theorem der Vertauschbarkeit zweier beliebiger Translationen $\overrightarrow{PQ}$ und $\overrightarrow{RS}$.*

Beweis. Nach Axiom I ist immer $PQRPQR$, durch Zwischenschieben der Identität SS folgt $PQRSSPQR$, oder wenn wir nach Axiom I das Tripel SPQ durch QPS ersetzen: $PQRSQPSR$.

§ 6.

Das zweite Zusatzaxiom.

Wir fordern:

Axiom II. *Es gibt keine Reihen, die eine ungerade Anzahl von kleinen lateinischen Buchstaben — [wiederholt vorkommende Buchstaben mehrfach gezählt] — enthalten.*

In der gewöhnlichen Geometrie haben wir in Axiom II einfach den Satz, daß eine Spiegelungsreihe mit einer ungeraden Anzahl von Geradenspiegelungen eine Umlegung ist, also nicht die Identität sein kann.

Wenn wir nach § 3 für die Tatsache, daß das Produkt PQR seinerseits wieder ein Punkt (eine Punktspiegelung) ist, auch die Bezeichnung verwenden: Es gilt $[PQR]$, so haben wir nach Axiom I und II den

Satz 9. Es gilt für drei beliebige Punkte immer $[PQR]$.

Nach Axiom I könnte höchstens PQR noch eine Gerade sein $PQR = a$. Das würde aber die nach Axiom II verbotene $PQRa$ zur Folge haben. In diesem Satz liegt die geometrische Tatsache: *Zu drei in bestimmter Reihenfolge genommenen Punkten P, Q, R gibt es stets genau einen vierten PQR], der mit ihnen im Parallelogramm liegt.*

Satz 10. *Für das Senkrechtstehen zweier Geraden ist notwendig und hinreichend das Bestehen von $[ab]$. Der Punkt $[ab]$ ist Schnittpunkt der beiden Geraden.*

Beweis. Weil für senkrechte Gerade ausdrücklich $(ab)^2$, aber nicht ab gefordert wird, könnte nur noch ab gleich einer Geraden c sein, was abc zur Folge hätte, also nach Axiom II unmöglich wäre. Daß $[ab]$ Schnittpunkt ist, folgt aus $aabaab = (a[ab])^2$ und $babbab = (b[ab])^2$.

Satz 11. $(aP)^2 + (bP)^2 + [ab] \rightarrow abP$.

Beweis. Die Reihe $ababPP$ ist wegen $(ab)^2$ trivialerweise richtig. Wegen $(aP)^2$ und $(bP)^2$ kann man nun P mit a und b beliebig vertauschen. Daher ergibt $ababPP$ auch $abPabP$. Das involutorische Element abP, das nach Axiom II keine Gerade sein darf, kann nun auch kein Punkt sein. Denn es wären dann $[ab]$, P und $[abP]$ drei Punkte, man könnte die Reihe $abPabP$ also auch als Reihe $[ab]P[abP]$ dreier Punkte schreiben, was dem Axiom I widerspricht. Also bleibt für das Element abP nur noch die Möglichkeit der Identität, und das war zu beweisen. Wegen $P = ab$ ist somit jeder Punkt P, für den Satz 11 gilt, gleich dem Punkt $[ab]$, und das ergibt nach Satz 10: *Zwei senkrechte Gerade haben genau einen Schnittpunkt.*

Satz 12. $[ab] + [bc] + [ac]$ können nicht gleichzeitig bestehen.

In Worten: *Drei paarweise senkrechte Gerade sind unmöglich.*

Beweis. Die trivialerweise richtige Reihe $abbcca$ könnte man auch als Reihe $[ab][bc][ca]$ dreier Punkte schreiben, was Axiom I widerspricht.

Satz 13. $[as] + [bs] + [cs] \rightarrow (abc)^2$.

Drei zu einer vierten Geraden senkrechte Gerade sind involutorisch in einem Büschel gelegen).

Beweis wurde in der Einleitung unter (21) gegeben.

Schreibt man die Tatsache, daß ein Produkt abc oder ein Produkt aP einer Geraden gleich ist, nach § 3 auch als $\langle abc \rangle$ und $\langle aP \rangle$, so ergibt sich

Satz 14. $(abc)^2$ ist gleichbedeutend mit $\langle abc\rangle$.

Satz 15. $(aP)^2$ ist gleichbedeutend mit $\langle aP\rangle$.

Denn abc und aP dürfen nach Axiom II weder die Identität noch Punkt sein. Die Gerade $\langle aP\rangle$ in Satz 15 ist, da $aaPaaP = a\langle aP\rangle a\langle aP\rangle$ gilt, und $a\langle aP\rangle$ nicht gilt, da man also $[a\langle aP\rangle]$ schreiben kann, zu a senkrecht und geht wegen $(\langle aP\rangle P)^2$ durch P hindurch.

Satz 16. *Reihen, die nur Punkte in ungerader Anzahl enthalten, sind unmöglich.*

Beweis. Eine Reihe $PQRST$ von fünf Punkten kann man nach Axiom I als Reihe $[PQR]ST$ von drei Punkten schreiben, ihre Gültigkeit widerspricht also Axiom I. Eine Reihe von $2n+1$ Punkten $PP'P''P'''$ $P^{(2n)}$ kann man in der Form

$$\Big[\ldots\Big|\big[\big[[PP'P'']P'''P''''\big]\ldots\big]\ldots\Big|\ldots P^{2n}\Big]$$

schreiben, ihre Gültigkeit widerspricht somit gleichfalls dem Axiom I.

Satz 17. $(PQ)^2 \rightarrow PQ$, also $P = Q$.

Beweis. Das involutorische Element PQ ist nach Axiom II keine Gerade, könnte also außer der Identität höchstens noch Punkt sein, dann ließe sich $PQPQ$ aber als $[PQ]PQ$ schreiben im Widerspruch zu Axiom I.

Satz 18. $(aP)^4 \rightarrow (aP)^2 \rightarrow \langle aP\rangle$.

Beweis. Wegen $(aP)^4$ ist $(aP)^2$ involutorisch und kann nach Axiom II keine Gerade, sondern nur die Identität oder ein Punkt sein. Letzteres ist aber unmöglich, denn es ließe sich dann $aPaPaPaP$ in der Form $[aPaP][aPa]P$ schreiben im Widerspruch zu Axiom I. Aus der allein bleibenden Möglichkeit $(aP)^2$ folgt nach Satz 15 $\langle aP\rangle$.

Satz 19. $PRQR + PSQS \rightarrow RS$.

Beweis. $P = RQR$ in die zweite Voraussetzung eingesetzt ergibt $RQRSQS$, nach Axiom I kann man dafür schreiben, indem man statt QRS das Tripel SQR setzt: $RSRQQS = RSRS$, woraus sich nach Satz 17 dann das gewünschte RS ergibt. Das heißt aber $R = S$:

In Worten: *Ein Punktepaar hat höchstens einen Mittelpunkt.*

Satz 20. $[abcdZ] \rightarrow [abcdR]$ für jeden Punkt R.

Beweis. Nach Axiom I oder Satz 9 gilt für die drei Punkte $[abcdZ]$ Z und R, wo R ganz willkürlich ist: $[[abcdZ]ZR]$ oder, da man nach § 3 die inneren eckigen Klammern weglassen darf, $[abcdZZR]$ oder $[abcdR]$, w. z. b. w.

Wenn $[abcdZ]$ für einen Punkt Z gilt, so gilt die Reihe also gleich für jeden beliebigen Punkt Z, $[abcdZ]$ ist also eine Beziehung zwischen

n Geraden a, b, c und d, die mit einem ganz beliebigen Hilfspunkt schrieben werden kann. Im folgenden werden wir für solche ganz liebigen Hilfspunkte immer den Buchstaben Z reservieren. In Übereinmmung mit Nr. 17 der Tabelle des § 2 definieren wir jetzt für unsere uppengeometrie: *Zwei je in bestimmter Reihenfolge genommene Geradenare a, b und d, c heißen Paare mit (gleichsinnig) kongruenten Winkeln, nn $[abcdZ]$ mit einem beliebigen und nach Satz* 20 *dann mit jedem iebigen Hilfspunkt Z gültig ist.* Diese Definition braucht zunächst gar chts zu tun zu haben mit den in Nr. 1, 4 und 9 des § 4 gegebenen Defiionen „senkrecht", „Mittellinie (Winkelhalbierende)" und „Schneiden ter der Periode $2n$". Wir können aber beweisen:

Satz 21. $[ab] + [cd] \rightarrow [abcdZ]$.

Beweis. Für die drei Punkte $[ab], [cd]$ und Z kann man nach tz 9 schreiben $[[ab][cd]Z]$ oder $[abcdZ]$.

In Worten: *Senkrechte Geradenpaare sind Paare mit kongruenten nkeln.*

Satz 22. $[ab] + [abcdZ] \rightarrow [cd]$.

Beweis. Für die drei Punkte $[ba] = [ab]$, $[abcdZ]$ und Z können r wieder ansetzen: $[[ba][abcdZ]Z] = [baabcdZZ] = [cd]$, w. z. b. w.

Haben ein Geradenpaar c, d und ein zweites senkrechtes Paar konuente Winkel, so ist das Paar b, c selbst senkrecht.

Satz 23. *Ist das Geradentripel a, b, c involutorisch, so sind alle vier ipel aus den vier Geraden a, b, c und $\langle abc\rangle$ involutorisch, und die are a, b und $\langle abc\rangle$, c haben gleichsinnig kongruente Winkel.*

Beweis. Die erste Behauptung ergibt sich aus der trivialen Richtigit von $(ba\langle abc\rangle)^2$, sowie $(cb\langle abc\rangle)^2$, sowie $(ca\langle abc\rangle)^2$, die zweite gt aus $[abc\langle abc\rangle Z] = [(abc)^2 Z]$, was wegen $(abc)^2$ erfüllt ist.

Satz 24. $adbd \rightarrow [adbdZ]$.

. Beweis trivial.

In Worten: *Ist d Mittellinie von a und b, so haben die Paare ad d db gleichsinnig kongruente Winkel.*

Als Spezialfall erhalten wir aus Satz 20 für $c = d$:

Satz 25. $[abZ] \rightarrow [abR]$ für jeden Punkt R.

In Übereinstimmung mit Nr. 11 der Tabelle des § 2 können wir daher zt definieren: *Zwei Gerade heißen parallel, wenn $[abZ]$ gilt für einen liebigen und nach Satz* 25 *dann für jeden Hilfspunkt Z.* Da $[abZ]$ $a = b$ erfüllt ist, sind dann auch zwei identische Gerade parallel. Zu esem Begriff ist zu bemerken, daß zunächst nichts darüber ausgemacht , ob zwei parallele Gerade Schnittpunkte haben oder nicht. Auch hat

dieser Begriff parallel zunächst nichts mit dem Begriff involutorisch fü ein Geradentripel zu tun. Wir können aber beweisen:

Satz 26. $\langle Pb\rangle + \langle Pc\rangle + [bcZ] \rightarrow bc$ oder $b = c$.

In Worten: *Zwei verschiedene parallele Gerade b und c haben keinen Schnittpunkt.*

Beweis. Aus $[bcZ]$ folgt nach Satz 20 $[bcP]$ oder $bcPbcP$. D man wegen der ersten beiden Voraussetzungen P mit b und c vertauschen kann, darf man dafür schreiben $bcbcPP$ oder $bcbc$. Wir haben nach Axiom II nur noch nachzuweisen, daß bc kein Punkt sein darf, da $bc =$ die zu beweisende Behauptung ist. Wir weisen nach, daß $[bc]$ und $[bcZ$ (dritte Voraussetzung) nicht zugleich bestehen können. Es ließe sich dann nämlich die richtige Reihe $cbbcZZ$ in der Form $[cb][bcZ]$ schreiben, was dem Axiom I widerspricht.

Wir betonen noch, daß durch unsern Satz 26 noch nichts darüber ausgemacht ist, daß nichtparallele Gerade, also solche, für die bei keinem die Reihe $[bcZ]$ gilt, einen Schnittpunkt besitzen.

Den Schluß des Beweises von Satz 26 formulieren wir als

Satz 27. $[bc] + [bcZ]$ dürfen nicht gleichzeitig gelten.

In Worten: *Senkrechte Gerade sind nie parallel.*

Satz 28. $[abZ] + [acT] \rightarrow [bcS]$.

In Worten: *Sind zwei Gerade einer dritten parallel, so sind sie untereinander parallel.*

Beweis. Es läßt sich ansetzen: $[T[Tca][abZ] = [TTcaabZ]$ oder $[cbZ]$, was aber nach § 1 (24) mit $[bcZ]$ und nach Satz 25 mit $[bcS]$ gleichbedeutend ist.

Satz 29. $[abZ] + [acT] + \langle Pb\rangle + \langle Pc\rangle \rightarrow bc$.

In Worten: *Ist eine Grade a und ein Punkt P gegeben, so gibt es durch P höchstens eine Parallele zu a.* (Über die Existenz einer solchen Parallelen ist nichts ausgemacht.)

Beweis. Aus Satz 28 folgt mittels der ersten beiden Behauptungen $[bcZ]$, hieraus und aus den letzten beiden Behauptungen dann nach Satz 26 daß $b = c$ ist.

Satz 30. $[abcdZ] \rightarrow [cdabZ]$.

Dies ist der noch nachzutragende Satz, daß die Beziehung der Winkelkongruenz zweier Geradenpaare reziprok ist.

Beweis. $[abZba]$ ist der mit ab transformierte Punkt Z, ihn kann man nach Satz 20 statt Z in $[abcdZ]$ einsetzen und erhält $[abcd[abZba]$. Vorn und hinten kann man a und b wegkürzen, und es bleibt das gewünschte $[cdabZ]$.

Satz 31. $[abcdZ] \rightarrow [dabcZ]$. [15])

Dies ist, wie die Fig. 10 zeigt, in der gewöhnlichen Geometrie der Satz: *etragen in einem Viereck zwei gegenüber-egende Winkel zwei Rechte, so tun es auch e andern beiden Winkel.*

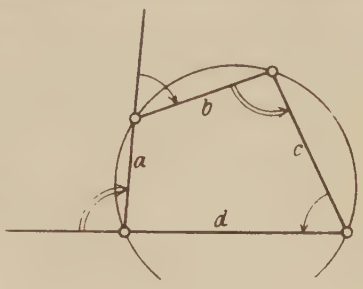


Fig. 10.

Beweis. Wir setzen in $[abcdZ]$ att Z den Punkt $[dZd]$ ein und erhalten $bcd[dZd]] = [abcZd]$ oder das gewünschte $[dabcZ]$.

Satz 32.

$[abcdZ] + [dcfgZ] \rightarrow [abfgZ]$.

Dieser Satz besagt die *Transitivität der Beziehung der Winkelkongruenz* für die Geradenpaare a, b und d, c und g, f.

Beweis. Wir können ansetzen $[[abcdZ]Z[dcfgZ]]$ oder $bcdZZdcfgZ]$ oder $[abfgZ]$, wie verlangt.

Satz 33. $[dabcZ] + [gdcfZ] \rightarrow [gabfZ]$.

Diesem Satz entspricht in der gewöhnlichen Geometrie die Aussage: *nd in zwei Dreiecken zwei Paare tsprechender Winkel gleichsinnig ngruent, so ist es auch das dritte ur.* Dies lehrt die Fig. 11. In serer Gruppengeometrie können wir ofern noch nicht von einem Dreiprechen, als wir noch nicht ssen, ob zwei nichtparallele Gerade einen Schnittpunkt haben.

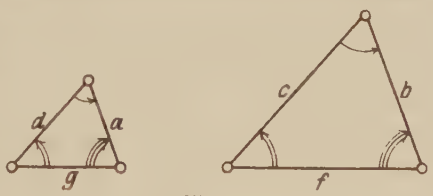


Fig. 11.

Beweis. Die beiden Voraussetzungen können wir nach Satz 31 als $bcdZ]$ und $[dcfgZ]$ schreiben, woraus nach Satz 32 folgt: $[abfgZ]$ er nach Satz 31 das verlangte $[gabfZ]$.

Setzt man in Satz 32 $b = c = f$, so erhält man als Spezialfall den Satz 28.

Satz 34. $[abZ] + [acZ] \rightarrow (abc)^2 \rightarrow \langle abc \rangle$.

In Worten: *Drei parallele Gerade sind involutorisch.*

Beweis. Aus $[acZ]$ folgt $[caZ]$, und man kann dann nach Axiom I setzen: $[abZ]Z[caZ][abZ]Z[caZ]$. Ersetzt man hierin die Darstellung vierten der sechs Punkte durch die gleichwertige $[Zba]$ und die des sechsten rch $[Zac]$, so hat man $abZZcaZZbaZZac$ oder $abcabc$, w. z. b. w.

[15]) Aus Satz 30 und Satz 31 zusammen folgt, daß man in $[abcdZ]$ mit a, b, c d alle acht Permutationen vornehmen kann, welche die Paare $\{ac\}$ und $\{bd\}$ rennt lassen.

Satz 35. $[abZ]+[ac]\rightarrow[bc]$.

In Worten: *Ist c zu a senkrecht, so auch zu jeder Parallelen von a.*

Beweis durch den Ansatz $[[ca][abZ]Z]$.

Satz 36. $[as]+[bs]\rightarrow[abZ]$.

In Worten: *Sind zwei Gerade zu einer dritten senkrecht, so sind sie parallel.*

Beweis durch den Ansatz $[[as][sb]Z]$.

Satz 37. $[ac]+[bd]+[abZ]\rightarrow[cdZ]$.

Dieser Satz sagt für die gewöhnliche Geometrie im Zusammenhang mit Satz 35 *die Existenz von Rechtecken, d. h. Vierecken mit lauter rechten Winkeln und zwei Paaren paralleler Seiten aus.*

Beweis durch den Ansatz $[[ca][aZb][bd]]$.

Satz 38. $\langle aP\rangle+\langle aQ\rangle+PRQR\rightarrow\langle aR\rangle$.

In Worten: *Der Mittelpunkt eines Punktepaares befindet sich mit einer Verbindungsgeraden desselben stets in vereinigter Lage.*

Beweis. Setzt man das sich aus der dritten Voraussetzung ergebende $P=RQR$ in $\langle aP\rangle$ oder $aPaP$ ein, so erhält man $aRQRaRQR$. Das Q an der dritten Stelle kann man hier wegen $\langle aQ\rangle$ durch aQa ersetzen und hat nach Herübernehmen des letzten R nach vorn: $RaRaQaRaRQ$. Hier kann man nach Axiom I statt der ersten drei Punkte $R[aRa]Q$ auch setzen $Q[aRa]R$ und erhält: $QaRaRaRaRQ$ oder $(aR)^4$ oder nach Satz 18 $\langle aR\rangle$, w. z. b. w.

Zum Schluß beweisen wir noch einen rein formalen Hilfssatz, den wir später brauchen:

Satz 39. $[pa]+\langle pQ\rangle+\langle pR\rangle\rightarrow(aRQ)^2\rightarrow\langle aRQ\rangle$.

Beweis. Wegen $\langle pQ\rangle$ ist auch $pQpQaa$ eine richtige Reihe. Hieraus ergibt sich $p=QaapQ$. Setzen wir das in $\langle pR\rangle$ oder $RpRp$ an zweite Stelle ein, so erhält man $RQaapQRp$. Hierin können wir nach Axiom I die Reihenfolge der drei Punkte $[ap]$, Q und R umkehren und erhalten: $RQaRQapp$ oder $(RQa)^2$ oder $(aRQ)^2$, w. z. b. w.

§ 7.

Das dritte Zusatzaxiom.

Axiom III. *Für drei beliebige Gerade a, b, c und einen beliebigen Punkt R gilt stets $((abc)^2R)^2$, aber nie $(abc)^2R$.*

In Rücksicht auf Axiom II ergibt sich aus diesem Axiom der

Satz 40. *Es gilt stets* $[(abc)^2R]$.

Die Richtigkeit des Axioms III in der gewöhnlichen Geometrie ergibt ch erstens daraus, daß $((abc)^2R)^2$ die im § 2 (27) schon abgeleitete lentische Reihe ist und zweitens daraus, daß aus $(abc)^2R = 1$ folgen ürde $(abc)^2 = R$. Nun haben wir aber im § 2 schon darauf hinge- iesen, daß das Quadrat einer Umlegung im besonderen also auch das on abc stets eine Translation ist. Die Translation $(abc)^2$ kann aber nie eich der Punktspiegelung R sein. Damit ist die Unmöglichkeit von $abc)^2R$ dargetan.

Wir beweisen jetzt mit den Axiomen I, II und III:

Satz 41. $(abc)^4 \rightarrow \langle abc\rangle$.

Beweis. Das nach Voraussetzung involutorische Element $(abc)^2$ ann nach Axiom II nur die Identität, wie es Satz 41 wegen $(abc)^2 = \langle abc\rangle$ ehauptet, oder eine Punktspiegelung sein. Letzteres haben wir ad absurdum ı führen. Für die drei Punkte $[(abc)^2]$, R und S, wo S ganz beliebig, önnten wir in diesem Falle nach Axiom I ansetzen: $[[(abc)^2]RS]$. Hierfür önnen wir aber auch schreiben $[(abc)^2]RS[(abc)^2]RS$. Da nun nach xiom III oder Satz 40 auch $(abc)^2R$ ein Punkt ist, können wir statt essen auch schreiben $[(abc)^2RS][(abc)^2R]S$. Wir hätten dann eine eihe dreier Punkte, was im Widerspruch zu Axiom I steht.

Satz 42. $abcRcbaR \rightarrow \langle abc\rangle$.

Beweis. Nach Voraussetzung gilt $abcR = Rabc$. Es ist also R it abc vertauschbar. Nach Axiom III gilt stets $(abc)^2R\ (abc)^2R$. ertauscht man nun darin das erste R zweimal hintereinander mit den eiden davorstehenden abc, so erhält man $R(abc)^4R$ oder $(abc)^4$. Daraus gibt sich aber nach Satz 41 das gewünschte $\langle abc\rangle$.

Da man die Voraussetzung des Satzes als $abcRcba = R$ schreiben ann und da $abcRcba$ der mit abc transformierte Punkt R ist, besagt er Satz 42 in der gewöhnlichen Geometrie, *daß eine aus drei Geraden- piegelungen erzeugte Umlegung, die einen Punkt fest läßt, selbst eine eradenspiegelung ist.*

Satz 43. $\langle Ra\rangle + \langle Rb\rangle + \langle Rc\rangle \rightarrow \langle abc\rangle$.

In Worten: *Drei Gerade, die durch denselben Punkt gehen, sind in- olutorisch (gehören demselben Büschel an).* Jetzt rechtfertigt sich also nsere Bezeichnung „demselben Büschel angehörend" insofern, als nach tz 34 und 43 sowohl drei parallele als auch drei durch denselben Punkt hende Gerade involutorisch sind. Allerdings ist nicht bewiesen, daß drei volutorische Gerade entweder parallel sind oder durch einen Punkt gehen.

Beweis. In der ersten Voraussetzung $RaRa$ setzen wir für das ste R den aus der zweiten folgenden Ausdruck $R = bRb$ ein und er- lten $bRbaRa$. Hier ersetzen wir wieder das erste R durch den aus

der dritten Voraussetzung folgenden Ausdruck $R = cRc$ und erhalten $bcRcbaRa$ oder $abcRcbaR$, daraus ergibt sich nach Satz 42 aber das gewünschte $\langle abc\rangle$.

Satz 44. $PbQb + QaRa + RcPc \to \langle abc\rangle$.

Dies ist in der gewöhnlichen Geometrie der *Satz, daß die drei Mittellote a, b, c eines Dreiecks mit den Ecken P, Q, R durch einen Punkt gehen.* Für unsere Gruppengeometrie haben wir vorläufig nur das Skelett dieses Satzes, weil wir noch nicht wissen, ob das Mittellot eines Punktepaares durch dessen Mittelpunkt geht und auf einer Verbindungsgeraden desselben senkrecht steht. Ebenso wissen wir ja noch nicht, daß für die drei als involutorisch nachgewiesenen Mittellote ein gemeinsamer Schnittpunkt existiert. *Immerhin werden wir in diesem Satze den eigentlichen Kern zum Mittellotsatz zu erblicken haben.*

Beweis. Aus der zweiten und dritten Voraussetzung folgt $Q = aRa$ und $P = cRc$; in die erste Voraussetzung eingesetzt ergibt das $cRcbaRab$ oder $abcRcbaR$, woraus nach Satz 42 die Behauptung folgt.

Wir können jetzt den Satz vom Schnittpunkt der drei Höhen eines Dreiecks fast vollständig beweisen. Wenn wir ganz ausführlich vorgehen, so ist ein Dreieck ein System von drei Punkten P, Q, R und drei Geraden p, q, r, für die gilt:

$$(*) \qquad \langle Pq\rangle + \langle Pr\rangle + \langle Qp\rangle + \langle Qr\rangle + \langle Rp\rangle + \langle Rq\rangle.$$

Sind a, b und c jetzt die drei Höhen dieses Dreiecks, so muß für sie gelten:

$$(**) \qquad \langle Pa\rangle + \langle Qb\rangle + \langle Rc\rangle + [pa] + [qb] + [rc].$$

Wir sprechen jetzt den Höhensatz in der Form aus:

Satz 45. Aus den zwölf in (*) und (**) gemachten Voraussetzungen folgt $\langle abc\rangle$.

Die einzige Unvollständigkeit beruht noch darin, daß wir nicht wissen, ob die drei als involutorisch nachgewiesenen Höhen a, b, c einen gemeinsamen Punkt haben.

Beweis. Durch dreimalige Anwendung des Satzes 39 ergibt sich zunächst:

$$[pa] + \langle pQ\rangle + \langle pR\rangle \to \langle aRQ\rangle,$$
$$[qb] + \langle qR\rangle + \langle qP\rangle \to \langle bPR\rangle,$$
$$[rc] + \langle rP\rangle + \langle rQ\rangle \to \langle cQP\rangle.$$

Es ist somit jetzt gültig

$$(Pa)^2 + (Qb)^2 + (Rc)^2 + (RQa)^2 + (PRb)^2 + (QPc)^2.$$

Nun wenden wir den Satz 4 dreimal an:

$$(Pa)^2+(RQa)^2 \to QRPaPQRa,$$
$$(Qb)^2+(PRb)^2 \to RPQbQRPb,$$
$$(Rc)^2+(QPc)^2 \to PQRcRPQc.$$

Fassen wir nun die Tripel $[RPQ]$, $[QRP]$ und $[PQR]$ zu Punkten zusammen, so können wir den Satz 44 anwenden:

$$[RPQ]b[QRP]b+[QRP]a[PQR]a+[PQR]c[RPQ]c \to \langle abc\rangle,$$

womit auch Satz 45 bewiesen ist.

Der Beweis entspricht dem der gewöhnlichen Geometrie durch Zurückführung auf den Mittellotsatz: Zieht man durch die drei Ecken eines Dreiecks die Parallelen zu den Gegenseiten, so entsteht ein neues, dessen Mittellote mit den Höhen des alten übereinstimmen. In der gewöhnlichen Geometrie sind die Punkte $PQR = RQP$, $QRP = PRQ$ und $RPQ = QPR$ die Ecken dieses neuen Dreiecks (Fig. 12).

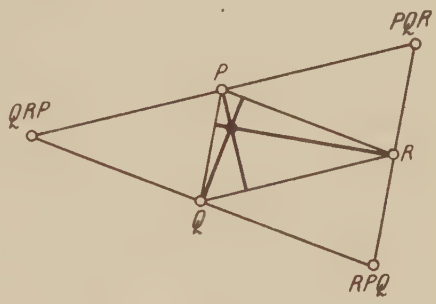


Fig. 12.

Satz 46. *Für sechs Gerade a, b, c, d, f, g gilt stets die identische Reihe:*

$$(abc)^2(dfg)^2(cba)^2(gfd)^2.$$

Beweis. Nach Axiom III sind $[(abc)^2R]$ und $[R(dfg)^2]$ Punkte. Nach Axiom I kann man also setzen

$$[(abc)^2R][R(dfg)^2]R[(abc)^2R][R(dfg)^2]R.$$

Hier darf man das zweite der $[(abc)^2R]$ durch $[R(cba)^2]$ ersetzen und das zweite der $[R(dfg)^2]$ durch $[(gfd)^2R]$. Man erhält so:

$$[(abc)^2R][R(dfg)^2]R[R(cba)^2][(gfd)^2R]R.$$

Nach Weglassen sämtlicher Klammern und Wegkürzung der drei Produkte RR folgt dann das gewünschte Resultat. Wenn wir $d=b$ sowie $f=c$ sowie $g=a$ setzen, so ergibt sich hier auch die identische Reihe (29) des § 2 als richtig.

Satz 47. $PRQR+PcQc \to \langle Rc\rangle$.

In Worten: *Der Mittelpunkt eines Punktepaares liegt notwendig auf einem Mittellot desselben.*

Beweis. Setzt man $P = RQR$, was aus der ersten Voraussetzung folgt, in die zweite ein, so hat man $RQRcQc$ oder $cRQ = QcR$; es ist also Q mit cR vertauschbar. Berücksichtigt man diese Tatsache in der sicher immer richtigen Reihe (Axiom I): $[cRc]RQ[cRc]RQ$, so erhält man $(cR)^4$, woraus nach Satz 18 die Behauptung $\langle Rc\rangle$ folgt. [Dieser Satz benutzt nur die Axiome I und II.]

§ 8.

Das vierte Zusatzaxiom.

Axiom IV. $abPQbaQP \rightarrow \left\{ \begin{array}{ll} \textit{entweder} & P = Q \\ \textit{oder} & (abP)^2 \end{array} \right\}$.

Dieses Axiom ist insofern von anderer Art als die ersten drei I, II, III, als hier nicht wie dort gewisse Reihen, die beliebige Punkte und Gerade enthalten, einfach als identisch richtig oder niemals gültig gefordert werden, sondern das Axiom IV hat die konditionale Form: Wenn die Reihe $abPQbaQP$ gilt, so soll notwendig gleichzeitig auch noch eine weitere Reihe erfüllt sein, nämlich entweder PQ (oder $P = Q$), oder wenn nicht diese, dann die Reihe $(abP)^2$. Statt der Eventualität: Entweder $P = Q$ oder $(abP)^2$ können wir auch die andere setzen: Entweder $P = Q$ oder $[abP]$. Denn wenn nach Axiom IV aus $abPQbaQP$ folgen würde abP, so wäre $ab = ba = P$, in $abPQbaQP$ eingesetzt ergäbe das $PPQPQP$ oder $QPQP$ oder nach Satz 17 weiter PQ oder $P = Q$. Es würde sich der zweite Fall für abP also auf den ersten $P = Q$ reduzieren. Daher können wir die zweite Eventualität $(abP)^2$ nach Axiom II auf $[abP]$ beschränken. Nach Satz 17 können wir somit aus Axiom IV die Aussage gewinnen:

Satz 48. $abPQbaQP \rightarrow \left\{ \begin{array}{ll} \text{entweder} & PQ \\ \text{oder} & [abP] \end{array} \right\}$.

In Worten: Sind P und Q zwei beliebige, aber verschiedene Punkte und a und b zwei beliebige Gerade, so zieht die Reihe $abPQbaQP$ notwendig nach sich, daß $[abP]$ Punktspiegelung ist. *Nach der im* § 6 *gegebenen Definition für den Parallelismus kann eine Reihe* $abPQbaQP$ *bei verschiedenen Punkten* P *und* Q *also nur bestehen, wenn* a *und* b *parallel sind.* Daraus erkennen wir dann auch die Richtigkeit unseres Axioms IV für die gewöhnliche Geometrie: $abPQbaQP$ besagt wegen $abPQ = PQab$ nämlich, daß die kongruenten Abbildungen ab und PQ vertauschbar sind. Falls nicht $P = Q$ und ab dann trivialerweise mit der Identität PQ vertauschbar ist, muß PQ eine wirkliche Translation sein, und ab ist dann

1. bei sich schneidenden Geraden eine wirkliche Drehung, 2. bei parallelen Geraden eine wirkliche Translation und 3. bei zusammenfallenden Geraden die Identität. Nur in den beiden letzten Fällen ist ab aber mit der Translation PQ vertauschbar, und diese beiden Fälle sind gerade durch $[abP]$ (oder $[abZ]$ mit beliebigem Hilfspunkt) gekennzeichnet.

Satz 49. $\langle adc\rangle + [adZ] \rightarrow \left\{\begin{array}{ll}\text{entweder} & a = d\\ \text{oder} & [caZ]\end{array}\right\}$.

In Worten kann man diesen Satz im Zusammenhang mit Satz 28 in der Form aussprechen: *Sind von drei involutorischen Geraden zwei nicht zusammenfallende parallel, so sind alle drei Geraden parallel.*

Beweis. Aus $\langle adc\rangle$ oder $cdacda$ folgt, daß auch $caZZadacdaZZ$ eine richtige Reihe ist. Wegen $Zad = daZ$ können wir diese Reihe aber in der Form schreiben: $caZ[Zad]ac[Zad]Z$. Diese Reihe hat aber gerade die Form der in Axiom IV oder Satz 48 vorausgesetzten Reihe, wobei c und a an die Stelle von a und b sowie Z und $[Zad]$ an die Stelle von P und Q treten. Nach Satz 48 muß daher entweder $Z[Zad]$, das heißt aber $a = d$ gelten oder $[caZ]$, w. z. b. w.

Satz 50. $cPcbPb \rightarrow \left\{\begin{array}{ll}\text{entweder} & \langle Pc\rangle + \langle Pb\rangle\\ \text{oder} & b = c\end{array}\right\}$.

Da umgekehrt mit $\langle Pc\rangle$ und $\langle Pb\rangle$ auch $cPcbPb$ erfüllt ist, wie man aus der Vertauschbarkeit von P mit b und c sieht, *kann man die Aussage: P ist ein Schnittpunkt der beiden verschiedenen Geraden b und c, auch durch die einzige Reihe $cPcbPb$* (allerdings mit der Nebenbemerkung $c \neq b$) *ausdrücken.*

Beweis. Aus $cPcbPb$ folgt $cPcbP = b$ oder durch „Quadrieren“ $bb = 1 = cPcbPcPcbP$. Diese Reihe kann man aber auch in der Form $cbPcPcbPcP$ oder $cbPcPcbccPcP$ oder $cbP[cPc]bc[cPc]P$ schreiben. Diese Reihe hat wieder die im Satz 48 geforderte Form, also kann man schließen: Entweder $P = cPc$ oder $[cbP]$. Ersteres führt wegen der Voraussetzung $cPcbPb$ aber auf $PbPb$, und nach Satz 15 erhält man hieraus und aus $P = cPc$ dann $\langle Pc\rangle$ und $\langle Pb\rangle$, also die erste Eventualität von Satz 50. Letzteres aber hat $cbP = Pbc$ zur Folge, in die Voraussetzung $cPcbPb$ in der Mitte eingesetzt ergibt das aber $PPbcb$ oder $(cb)^2$. Nach Satz 27 kann $(bc)^2$ aber neben $[cbP]$ nur dann bestehen, wenn nicht $[cb]$, sondern $b = c$ gilt. Damit haben wir die zweite Eventualität des Satzes erhalten.

Satz 51. $PcPbcb \rightarrow \left\{\begin{array}{ll}\text{entweder} & \langle Pc\rangle + (bc)^2\\ \text{oder} & \langle Pb\rangle + [bcP]\end{array}\right\}$.

Beweis. Aus $PcPbcb$ erhält man $PcPbc = b$ und $bb = 1 = PcPbcPcPbc$. Für diese Reihe kann man aber auch schreiben: $bcPcPbcPcP$ oder $bcPcPccbcPcP$ oder $bcP[cPc]cb[cPc]P$. Wendet man hierauf Satz 48 an, so ergibt sich: Entweder $P = cPc$ oder $[bcP]$. Aus ersterem folgt aber wegen $PcP = c$ und der Voraussetzung $PcPbcb$ die Reihe $cbcb$, wir haben also den ersten der in Satz 51 gefolgerten Fälle. Für $[bcP]$ ist aber $Pbc = cbP$. In die Mitte von $PcPbcb$ eingesetzt ergibt das aber $PccbPb$ oder $PbPb$, so daß wir auf den zweiten der angekündigten Fälle kommen.

Satz 52. $\langle aP\rangle + \langle aQ\rangle + \langle bP\rangle + \langle bQ\rangle \rightarrow \left\{\begin{array}{ll}\text{entweder} & a = b\\ \text{oder} & P = Q\end{array}\right\}$.

In Worten: *Zwei verschiedene Gerade haben höchstens einen Schnittpunkt* und: *Zwei verschiedene Punkte haben höchstens eine Verbindungsgerade.*

Beweis. Aus $\langle aP\rangle$ oder $(aP)^2$ und $(aQ)^2$ folgt $a = PaP = QaQ$, also $PaPQaQ$. Ebenso folgt aus $\langle bP\rangle$ und $\langle bQ\rangle$ die Reihe $PbPQbQ$. Aus der ersten der gewonnenen Reihen ergibt sich dann weiter $PQ = aPQa$, aus der zweiten $PQ = bPQb$; somit hat man $aPQa = bPQb$ oder $aPQabQPb$ oder $baPQabQP$, worauf man dann Satz 48 anwenden kann. Man erhält entweder $P = Q$ oder $[baP]$. Letzteres führt nach $baPbaP$ wegen der Vertauschbarkeit von P mit a und b aber auf $(ba)^2$. Nach Satz 27 kann $(ba)^2$ aber nur für $b = a$ mit $[baP]$ vereinbar sein. Damit haben wir die beiden in unserm Satz behaupteten Fälle.

Satz 53. $PcQc + PbQb \rightarrow \left\{\begin{array}{ll}\text{entweder} & P = Q\\ \text{oder} & b = c\end{array}\right\}$.

In Worten: *Zwei verschiedene Punkte haben höchstens ein Mittellot.*

Beweis. Aus den beiden Voraussetzungen ergibt sich $P = cQc = bQb$, also $cQcbQb$; nach Satz 50 folgt dann: entweder $\langle Qc\rangle$ und $\langle Qb\rangle$ oder $b = c$. Im ersten Fall könnte man aber wegen $cQc = Q$ aus der ersten Voraussetzung $P = Q$ ableiten, womit der Satz bewiesen ist.

Satz 54. $\langle aP\rangle + \langle bP\rangle + \langle abc\rangle \rightarrow \left\{\begin{array}{ll}\text{entweder} & a = b\\ \text{oder} & \langle cP\rangle\end{array}\right\}$.

In Worten: *Gehen von drei involutorischen Geraden zwei nicht zusammenfallende durch einen Punkt, so geht auch die dritte durch diesen Punkt.*

Beweis. Setzt man $P = aPa$ in die zweite Voraussetzung $bPbP$ an zweite Stelle ein, so hat man $baPabP$. Aus $\langle abc\rangle$ folgt nun $ba = cabc$. Setzt man das am Anfang für ba in $baPabP$ ein, so hat

nan $cabcPabP$ oder $cabcPccabP$ oder $\langle cab\rangle cPc\langle cab\rangle P$ oder $Pc\langle cab\rangle P\langle cab\rangle$. Hierauf kann man Satz 50 anwenden und bekommt: ntweder $\langle Pc\rangle$ (und $\langle Pcab\rangle$) oder $c = cab$, d. h. aber $a = b$, womit er Satz bewiesen ist.

Satz 55. $\langle aP\rangle + \langle aQ\rangle + PcQc \rightarrow \left\{\begin{array}{ll}\text{entweder} & P = Q\\ \text{oder} & [ac]\end{array}\right\}$.

n Worten: *Das Mittellot zweier verschiedener Punkte ist auf der Verindungsgeraden immer senkrecht.*

Beweis. Setzt man das aus der dritten Voraussetzung folgende $= cQc$ in die erste ein, so hat man $acQcacQc$; ersetzt man hierin as zweite Q nach der zweiten Voraussetzung durch aQa, so bekommt nan $acQcacaQac$ oder $cQc\langle aca\rangle Q\langle aca\rangle$. Hierauf kann man den atz 50 anwenden, das ergibt: entweder $\langle Qc\rangle$ (und $\langle Qaca\rangle$) oder $= aca$. Im ersten Fall liefert $Q = cQc$ in die Voraussetzung $PcQc$ ingesetzt $P = Q$. Im zweiten Fall könnte $a = c$ oder $[ac]$ gelten. Es ürde aber $c = a$ nach der zweiten der Voraussetzungen des Satzes 55 ieder auf $\langle Qc\rangle$ führen, und wie im ersten Fall erhielte man daraus $= Q$. Somit bleibt $[ac]$ übrig, und man kann also gerade den Satz 55 ussprechen.

Satz 56. $\langle aP\rangle + \langle bP\rangle + acbc \rightarrow \left\{\begin{array}{ll}\text{entweder} & a = b\\ \text{oder} & \langle Pc\rangle\end{array}\right\}$.

n Worten: *Schneiden sich zwei verschiedene Gerade in einem Punkte, so eht jede Mittellinie (Winkelhalbierende) des Geradenpaares durch diesen unkt hindurch.*

Beweis. $a = cbc$ in $(aP)^2$ eingesetzt ergibt $cbcPcbcP$; nach der weiten Voraussetzung kann man hier das letzte P durch bPb ersetzen, o daß man hat: $cbcPcbcbPb$ oder $cPc\langle bcb\rangle P\langle bcb\rangle$. Durch Anendung von Satz 50 erhält man: entweder $\langle Pc\rangle$ (und $\langle Pbcb\rangle$) oder $= bcb$. Letzteres liefert wegen $acbc$ aber $a = b$. Damit ist das Beauptete bewiesen.

Satz 57.

$$PQRS + \langle aP\rangle + \langle aQ\rangle + \langle bR\rangle + \langle bS\rangle \rightarrow \left\{\begin{array}{lll}\text{enweder} & P = Q \text{ und } R = S\\ \text{oder} & [abZ]\end{array}\right\}.$$

n Worten: *Wenn die vier Punkte P, Q, R und S im Parallelogramm iegen* und dabei die aufeinander folgenden Punkte P und Q ebenso wie $?$ und S nicht zusammenfallen, *dann sind eine Verbindungsgerade von P und Q und eine Verbindungsgerade von R und S stets parallel.* Hier ndet also die Bezeichnung „Im Parallelogramm liegen“ für die Quadrupel it $PQRS = 1$ erst ihre eigentliche Rechtfertigung.

Beweis. $P = QRS$ in $\langle aP\rangle$ eingesetzt ergibt $aQRSaQRS$ oder wegen der aus $\langle aQ\rangle$ folgenden Vertauschbarkeit von Q mit a: auch $aQRSQaRS$. Nach Axiom I kann man hier statt QRS das Tripel SRQ einsetzen, so daß man $aSRQQaRS$ oder $SR = aSRa$ bekommt. Da aus $\langle bR\rangle$ und $\langle bS\rangle$ aber weiter $SR = bSRb$ folgt, hat man $aSRa = bSRb$ oder $abSRbaRS$. Hieraus bekommt man mittels Satz 48: entweder $S = R$ und nach $PQRS$ dann auch $P = Q$, oder $[abZ]$, wie behauptet.

Satz 58. Falls zwei Gerade a und b einen Punkt gemein haben oder parallel sind, gilt:

$$\langle abc\rangle + \langle abd\rangle \rightarrow \left\{\begin{array}{ll} \text{entweder} & \langle bcd\rangle + \langle acd\rangle \\ \text{oder} & a = b \end{array}\right\}.$$

Dieser Satz ist nur eine Zusammenfassung schon gewonnener Sätze. Haben a und b nämlich den Punkt P gemein, so folgt für $a \neq b$ aus $\langle abc\rangle$ und $\langle abd\rangle$ mittels Satz 54, daß auch $\langle Pc\rangle$ und $\langle Pd\rangle$ gilt. Mit Satz 43 kann man daraus dann $\langle bcd\rangle$ und $\langle acd\rangle$ ableiten. Sind aber a und b parallel, aber nicht identisch, so folgt aus Satz 49 wegen $[abZ]$ und $\langle abc\rangle$ sowie $\langle abd\rangle$, daß auch $[acZ]$ und $[adZ]$ gilt. Dann ergibt aber Satz 34 die gewünschten Behauptungen.

Wir kommen jetzt zu dem *Beweis des Satzes vom Schnittpunkt der drei Winkelhalbierenden x, y und z eines Dreiecks mit den Seiten a, b und c.* Wir können den Satz nicht in der folgenden Art formulieren:

$$azbz + bxcx + cyay \rightarrow \langle xyz\rangle.$$

Denn in dieser Form ist der Satz ja in der gewöhnlichen Geometrie nicht richtig. Vielmehr gibt es durch jeden Eckpunkt des Dreiecks zwei Winkelhalbierende, eine innere und eine äußere, und nur, wenn wir geeignete Tripel wählen, schneiden sich diese in einem und demselben Punkt. Wir formulieren daher den Satz folgendermaßen:

Satz 59.

$$\left\{\begin{array}{c} azbz + aycy + \langle Py\rangle + \langle Pz\rangle + \langle xyz\rangle + \langle xbc\rangle \\ (abc)^2 \neq 1 \end{array}\right\} \rightarrow bxcx.$$

Die Einschränkung $(abc)^2 \neq 1$ machen wir, weil wir von einem Dreieck nur dann sprechen werden, wenn die drei Seiten nicht involutorisch sind. Wir nehmen nach $azbz$ und $aycy$ dann zwei Winkelhalbierende y und z des Dreiecks als vorliegend an und nehmen weiter an, daß sie durch einen gegebenen Schnittpunkt P hindurchgehen. Es wird dann weiter eine Gerade x angenommen, die sowohl mit y und z zusammen wie auch mit b und c zusammen involutorisch ist. Diese Bedingungen stehen an Stelle der Aussage: „x verbindet P mit dem Eckpunkt des Dreiecks, von dem die Seiten b und c auslaufen."

Beweis. Aus der ersten Voraussetzung ergibt sich $a = zbz$ und $c = yzbzy$. Setzt man diese Werte für a und c überall in Satz 59 ein, so bleibt der Satz zu beweisen nach:

$$(*) \qquad \left\{ \begin{array}{c} \langle Py\rangle + \langle Pz\rangle + \langle xyz\rangle + \langle xbyzbzy\rangle \\ (zbzbyzbzy)^2 \neq 1 \end{array} \right\} \rightarrow bxyzbzyx.$$

Den Beweis dieses Satzes (*) führen wir in folgenden Schritten:

α) Es ist $y \neq z$, da sonst die in (*) vorausgesetzte Ungleichung nicht bestehen könnte.

β) Nach (*) ist durch $\langle xyz\rangle$ eine Gerade

$$t = xyz = zyx$$

gegeben, und wegen der weiteren Voraussetzung $\langle xbyzbzy\rangle$ oder $\langle yzbzyxb\rangle$ oder $\langle xyzbzyxbx\rangle$ oder $\langle tbtbx\rangle$ ist auch durch

$$s = tbtbx$$

eine solche gegeben.

γ) Trivialerweise gelten jetzt die drei Reihen

$$\langle tbtbt\rangle \text{ und } \langle tbtbxxt\rangle \text{ und } \langle tbtbxxb\rangle,$$

die wir auch als

$$\langle syz\rangle \text{ und } \langle sxt\rangle \text{ und } \langle sxb\rangle$$

schreiben können.

δ) Weil nach α) $y \neq z$ ist, folgt nach Satz 54 aus $\langle Py\rangle$ und $\langle Pz\rangle$ und $\langle xyz\rangle$, daß auch $\langle Px\rangle$ gilt. Ebenso folgt $\langle Ps\rangle$ aus $\langle Py\rangle$ und $\langle Pz\rangle$ und [vgl. γ)] $\langle syz\rangle$.

ε) Wegen [vgl. δ)] $\langle Px\rangle$ und $\langle Ps\rangle$ kann man mittels Satz 58 aus den in γ) gewonnenen Reihen $\langle sxt\rangle$ und $\langle sxb\rangle$ folgern, daß

entweder I. $\langle xtb\rangle + \langle stb\rangle$

oder II. $s = x$

sein muß.

ζ) Die Eventualität II in ε) führt [vgl. β)] auf $btbt$ oder $bxyzbzyx$, also auf die Behauptung des Satzes (*).

ϑ) Es bleibt somit nur die Aufgabe, die Eventualität I in ε) ad absurdum zu führen. Aus $\langle xtb\rangle$ folgt nun nach β) $\langle xxyzb\rangle$ oder $\langle yzb\rangle$. Wegen $xyz = zyb$ kann man dann die in den Voraussetzungen von (*) vorkommende letzte Ungleichung als $(zbzzybbzy)^2 \neq 1$ oder $(yzbzy)^2 \neq 1$ schreiben, was auf einen Widerspruch führt, denn es ist $(yzbzy)^2$ trivialerweise $= 1$.

§ 9.

Das fünfte Zusatzaxiom. Schlußbetrachtungen.

Axiom V. $(PQ)^3 \rightarrow P = Q$.

Dieses Axiom ist in der gewöhnlichen Geometrie sicher erfüllt, denn es besagt, daß eine dreimal hintereinander angewendete Translation um eine von 0 verschiedene Strecke nie gleich der Identität sein kann.

Satz 60. $(aP)^6 \rightarrow \langle aP\rangle$.

Beweis durch $([aPa]P)^3$ und Anwendung von Axiom V sowie Satz 15.

Satz 61. $\langle aP\rangle + \langle aQ\rangle + PRQRQR \rightarrow \langle aR\rangle$.

In Worten: *Ein* 2:1*-Teilpunkt eines Punktepaares* P, Q *liegt stets auf einer Verbindungsgeraden von* P *und* Q.

Beweis. Setzt man $P = RQRQR$ in $(aP)^2$ ein, so ergibt sich $aRQRQRaRQRQR$. Ersetzt man hierin den zweiten und dritten der vier Buchstaben Q nach $\langle aQ\rangle$ durch aQa, so hat man $aRaQaRQRaRaQaRQR$ oder, wenn man QR nach vorn nimmt: $\underline{QR[aRa]QaRQR[aRa]QaR}$. Ersetzt man hier die beiden unterstrichenen Punkttripel $QR[aRa]$ nach Axiom I durch $[aRa]RQ$, so bekommt man $aRaRQQaRaRaRQQaR$ oder $(aR)^6$, woraus nach Satz 60 das gewünschte $\langle aR\rangle$ folgt.

Satz 62. $PRQRQR + PSQSQS \rightarrow R = S$.

In Worten: *Ein geordnetes Punktepaar besitzt höchstens einen* 2:1*-Teilpunkt.*

Beweis. Aus beiden Voraussetzungen folgt durch Eliminieren von P die Reihe $\underline{RQRQRSQSQS}$. Stellt man nach Axiom I die unterstrichenen Tripel um, so erhält man $RSRSRQQSQQ$ oder $(RS)^3$, wonach mit Axiom V die Behauptung folgt.

Wir können jetzt die vier Sätze vom Schnittpunkt der Mittellinien, der Mittellote, der Höhen und der Winkelhalbierenden in dem folgenden Sinne als völlig bewiesen ansehen: Wir können uns jedesmal ein Dreieck mit den Eckpunkten P, Q und R und den Seitengeraden p, q und r gemäß den Relationen (*) des § 7 gegeben denken. Für Punkttripel P, Q, R, für die drei Verbindungsgerade p, q und r nicht existieren, interessieren wir uns nicht. Dann wissen wir jetzt nach Satz 52, daß p, q und r die einzigen möglichen Verbindungsgeraden der drei Paare der Dreiecksecken sind und ebenso P, Q und R die einzigen möglichen Schnittpunkte der drei Paare von Seitengeraden. Was den Satz über die Mittellinien (vgl. Satz 6) angeht, so wissen wir weiter aus Satz 19 und 38, daß es zu jedem Eckpunktepaar außer dem einen als gegeben angenommenen Mittelpunkt keinen weiteren gibt und daß der Mittelpunkt auf der Verbindungsgeraden des Eckpunktepaares liegt

Für Dreiecke, wo es zu mindestens einem Eckenpaar keinen Mittelpunkt gibt, interessieren wir uns nicht. Weiter interessieren wir uns dann nur für solche Dreiecke, bei denen zu jedem Eckpunkt und dem gegenüberliegenden „Seitenmittelpunkt" immer sowohl eine Verbindungsgerade (die Mittellinie), als auch ein 2:1-Teilpunkt und nach Satz 52 und 62 von jedem dann kein weiteres Exemplar existiert. Nach Satz 6 ist dann dieser 2:1-Teilpunkt alle drei Male derselbe, und die drei Mittellinien gehen nach Satz 61 durch ihn hindurch.

Was den Mittellotsatz (vgl. Satz 44) angeht, so wissen wir jetzt nach den Sätzen 47, 55 und 53, daß das Mittellot eines Eckpunktepaares des Dreiecks durch den Mittelpunkt desselben hindurchgeht und auf der Verbindungsgeraden senkrecht steht, falls eine solche vorhanden, wie wir natürlich annehmen. Ferner wissen wir, daß außer dem einen als gegeben angenommenen Mittellot des Eckpunktepaares kein weiteres vorhanden ist. Nach Satz 44 wissen wir dann, daß drei Mittellote eines Dreiecks involutorisch sind. Nehmen wir nun an, daß zwei der Mittellote einen (und nach Satz 52 dann nur einen) Schnittpunkt S gemeinsam haben, so geht nach Satz 54 dann auch das dritte Mittellot durch S hindurch. Die Existenz eines Schnittpunktes zweier der drei Mittellote müssen wir hier allerdings noch fordern.

Was den Höhensatz anlangt, so bemerken wir zunächst, *daß es zu einem Eckpunkt P und einer Gegenseite p höchstens eine Höhe gibt*, d. h. eine Gerade a, für die $\langle Pa\rangle$ und $[pa]$ gilt. Denn gäbe es zwei verschiedene a und b, so würde aus $\langle Pa\rangle + \langle Pb\rangle + [pa] + [pb]$ wegen Axiom I folgen: $[bp][pa]P]$ oder $[baP]$. Das steht aber nach Satz 26 zu $\langle Pa\rangle$ und $\langle Pb\rangle$ im Widerspruch. Nehmen wir nun drei Höhen als vorliegend an, so können wir nach Satz 45 beweisen, daß sie involutorisch sind. Nehmen wir dann wieder wie beim Mittellotsatz für zwei Lote hier für zwei Höhen einen gemeinsamen Schnittpunkt an, so folgt aus Satz 54, daß auch die dritte Höhe hindurchgeht. Zu dem Satz von den Winkelhalbierenden in der Fassung von Satz 59 haben wir nur als Ergänzung noch auf den Satz 56 hinzuweisen, wonach eine Winkelhalbierende eines Seitenpaares immer durch deren Schnittpunkt hindurchgeht. Ferner können wir beweisen, *daß zwei sich in einem Punkt P schneidende Gerade q und r stets entweder keine oder genau zwei senkrechte Mittellinien (Winkelhalbierende) besitzen*, was wir ausdrücken im

Satz 63. $$\left\{\begin{array}{c}\langle Pq\rangle + \langle Pr\rangle + qara + qbrb \\ a \neq b\end{array}\right\} \rightarrow [ab] + abP.$$

Beweis. Nach Satz 56 gilt auch $\langle Pa\rangle$ und $\langle Pb\rangle$. Nach Satz 43 ist also jedes Tripel aus dem Quadrupel q, r, a, b involutorisch. Es gilt also z. B. $abr = rba$. Nun folgt aus $qara$ und $qbrb$ durch Eliminieren von

q die Reihe $arabrb$. Ersetzt man darin das unterstrichene Tripel ab durch rba, so ergibt sich $abab$. Wegen $a \neq b$ gilt nun $[ab]$. Nach Satz 11 folgt aus $\langle Pa\rangle$, $\langle Pb\rangle$ und $[ab]$ dann abP, also $b = aP$, womit unser Satz 63 bewiesen ist. Man ersieht auch, daß es nicht mehr als zwei Mittellinien geben kann. Denn für eine dritte Mittellinie c mit $c \neq a$ könnte man ebenso wie für b die analoge Beziehung $c = aP$ ableiten, und es müßte $b = c$ sein. Gibt es nun umgekehrt schon eine Mittellinie a mit $qara$, so ist die zu a in P senkrechte Gerade $\langle aP\rangle$ immer eine zweite. Denn es gilt auch $q\langle aP\rangle r\langle aP\rangle$, was aus der Vertauschbarkeit von P mit r und a ohne weiteres ersichtlich. Damit ist dann auch die dem Satz 63 voraufgehende geometrische Behauptung bewiesen.

Wir kommen jetzt zu der *Unabhängigkeit der Axiome* I *bis* V. Wir haben fünf Unabhängigkeitsbeweise durch Pseudogruppengeometrien zu führen, indem wir von jedem Axiom nachweisen, daß es von den vier übrigen unabhängig ist. Von den Unabhängigkeitsbeweisen fehlt bislang noch der eine, nämlich der Nachweis, daß Axiom IV von I, II, III und V unabhängig ist. Immerhin scheint es äußerst unwahrscheinlich, daß sich IV aus den übrigen Axiomen beweisen läßt.

Die übrigen vier Beweise lauten so:

1. Unabhängigkeit von I. Man nehme die vollständige Permutationsgruppe von vier Elementen und nehme die Klasseneinteilung der involutorischen Elemente in Punkte und Geraden in der in der Einleitung angegebenen Weise vor. Die vollständige Untersuchung zeigt dann, daß in dieser Gruppengeometrie die Axiome II, III, IV und V erfüllt sind, während I nicht gilt, da für drei Punkte die Reihe PQR möglich ist[16]).

2. Unabhängigkeit von II. Man nehme die Gruppe der Punktspiegelungen und Translationen der gewöhnlichen Geometrie. Außer der Identität sind darin nur die Punktspiegelungen involutorisch. Man teile diese in die zwei Klassen ein: Spiegelungen an Punkten mit positiver kartesischer x-Koordinate nenne man Punkte, alle übrigen Punktspiegelungen Gerade. Dann sind I, III, IV und V erfüllt, II aber nicht.

3. Unabhängigkeit von III. Man nehme die gewöhnliche Gruppe aller gleichsinnig und ungleichsinnig flächentreuen Affinitäten der Ebene. Involutorisch sind außer der Identität nur die Punktspiegelungen, die man als „Punkte“, und die schiefen Spiegelungen an Geraden, die man als „Gerade“ betrachtet. Es gelten dann I, II, IV und V, aber nicht III.

---

[16]) Es wäre wünschenswert, einen Beweis zu finden für die Unabhängigkeit des ersten Teils von Axiom I, nämlich der Annahme, daß $(PQR)^2$ gilt, von den Axiomen II, III, IV, V und von der Aussage des Axioms I, daß PQR nie gilt. Ähnliches gilt für den ersten Teil von Axiom III.

4. Unabhängigkeit von V. Man nehme die Untergruppe aller derjenigen Permutationen von sechs Elementen $(abcdfg)$, bei denen die ersten drei Elemente unter sich und die letzten drei Elemente unter sich permutiert werden (also die Gruppe, die das direkte Produkt der vollständigen Permutationsgruppe von drei Elementen mit sich selbst darstellt.) Als Gerade nehme man die involutorischen Permutationen dieser Gruppe, bei denen nur zwei Elemente (entweder im ersten oder im letzten Tripel) vertauscht werden, alle übrigen aber fest bleiben. Als Punkte nehme man die übrigbleibenden von der Identität verschiedenen involutorischen Permutationen, bei denen in jedem Tripel zwei Elemente vertauscht werden. Es gelten in der zugehörigen Gruppengeometrie die Axiome I, II, III, IV, aber nicht V.

Wir geben jetzt eine kleine Zusammenstellung von Sätzen, die in der gewöhnlichen Geometrie richtig sind und die auf Grund von Axiom I bis V man beweisen bisher nicht gelungen ist. Sie können eventuell aus diesen Axiomen überhaupt nicht beweisbar sein:

Satz 64. $abxabxbaxbax + \{abxabx \neq 1\} \rightarrow [abZ]$.

Satz 65. $abxabxbaxbax + \{abxabx \neq 1\} \rightarrow \left\{\begin{matrix} abzabzbazbaz \\ \text{für jede Gerade } z \end{matrix}\right\}$.

Aus Satz 64 und 65 ersieht man, *daß man den Parallelismus zweier Geraden a und b statt mit einem beliebigen Hilfspunkt Z auch mit einer beliebigen Hilfsgeraden z erklären kann.*

Satz 66. $\langle abc\rangle + \langle abd\rangle \rightarrow \left\{\begin{matrix} \text{entweder } \langle bcd\rangle + \langle acd\rangle \\ \text{oder} \qquad a = b \end{matrix}\right\}$.

Dieser Satz konnte als Satz 58 nur unter einschränkenden Voraussetzungen bewiesen werden.

Satz 67. $azbz + aycy + bxcx + btct \rightarrow \left\{\begin{matrix} \text{entweder } \langle xyz\rangle \\ \text{oder} \qquad \langle tyz\rangle \end{matrix}\right\}$.

Dieser Satz ist eine gewisse Fassung des Satzes vom Schnittpunkt der Winkelhalbierenden, ohne daß für das Dreiseit a, b, c Eckpunkte vorausgesetzt werden.

Satz 68. $\langle pqc\rangle + \langle qra\rangle + \langle rpb\rangle + [cr] + [bq] + [ap] \rightarrow \langle abc\rangle$.

Dies ist der Höhensatz bei einem Dreiseit, für das keine Eckpunkte vorausgesetzt werden.

Satz 69. $abPQbaPQ \rightarrow \left\{\begin{matrix} \text{entweder } P = Q \\ \text{oder} \qquad [ab] \end{matrix}\right\}$.

Anhang.

Satz. *Eine aus ihren involutorischen Elementen erzeugbare Gruppe ist notwendig zu der Gruppe der kongruenten Abbildungen in der analytischen Geometrie eines geometrischen Körpers isomorph, wenn sich ihre von der Identität verschiedenen involutorischen Elemente in zwei zueinander fremde Teilmengen („Punkte" und „Geraden"), deren erste mindestens zwei Elemente enthalten muß, einteilen lassen, so daß folgendes gilt:*

A. *Es gibt keine Gruppenrelationen aus lauter involutorischen Elementen, bei denen Elemente der zweiten Teilmenge (Wiederholungen desselben Elements mitgezählt) in ungerader Anzahl vorkommen.*

B. *Für ein Element R der ersten und drei Elemente a, b und c der zweiten Teilmenge gilt stets $((abc)^2R)^2 = 1$, aber nie $(abc)^2R = 1$.*

C. *Für zwei verschiedene Elemente P und Q der ersten und für zwei Elemente a und b der zweiten Teilmenge, für die $abPQbaQP = 1$ ist, gilt notwendig gleichzeitig auch immer $(abP)^2 = 1$.*

D. *Zu zwei Elementen P und Q der ersten Teilmenge gibt es mindestens eins der zweiten, a, für das gleichzeitig $PaPa = 1$ und $QaQa = 1$ erfüllt ist*[17]).

E. *Zu zwei Elementen der zweiten Teilmenge a und b gibt es mindestens ein Element P der ersten, für das entweder gleichzeitig $PaPa = 1$ und $PbPb = 1$ oder für das $(abP)^2 = 1$ gilt.*

F. *Zu zwei Elementen a und b der zweiten Teilmenge gibt es mindestens ein weiteres c der zweiten Teilmenge, für das $acbc = 1$ gilt.*

Den genauen Beweis bringen wir in einer späteren Arbeit, hier machen wir nur die folgenden Andeutungen: A, B und C sind das zweite, dritte und vierte Zusatzaxiom der §§ 6 bis 8. Wesentlich ist, daß wir jetzt im Gegensatz zu unseren früheren Zusatzaxiomen neue Existenzannahmen machen, die über die in den Gruppenaxiomen enthaltenen Existenzannahmen hinausgehen. D verlangt die Existenz mindestens einer Verbindungsgeraden zweier Punkte, E verlangt die Existenz mindestens eines Schnittpunkts zweier Geraden a und b, für die mit keinem Punkt Z die Relation $(abZ)^3 = 1$ erfüllt ist. F fordert die Existenz mindestens einer Mittellinie eines Geradenpaares. Der Beweis verläuft nun in folgenden Schritten: Zuerst zeigt man, daß es zu zwei verschiedenen Punkten P

[17]) Statt D und E kann man auch die schwächeren Axiome nehmen:

D*. Zu P und Q gibt es mindestens ein a mit $PaPQaQ = 1$;

E*. Zu a und b gibt es mindestens ein P, für das entweder $aPabPb = 1$ oder $(abP)^2 = 1$ gilt.

Aus A, B, C, D*, E* und F lassen sich D und E ableiten.

nd Q stets mindestens eine Gerade c mit $PcQc$, *also mindestens ein Mittellot*, gibt. (Das $=1$ in den Gruppenrelationen lassen wir jetzt wieder eg und verwenden ganz die Bezeichnungen der vorangehenden Paraaphen.) Zu dem Zweck bemerken wir zuerst, daß es nach Axiom D ine Verbindungsgerade b von P und Q gibt: $(Pb)^2 + (Qb)^2$. Nach xiom A sind dann $\langle Pb\rangle$ und $\langle Qb\rangle$ Gerade. Nach Axiom F gibt es dann u Pb und Qb eine Mittellinie c mit $\langle Pb\rangle c\langle Qb\rangle c$. Wir wollen zeigen, aß dies c zugleich das gesuchte Mittellot von P und Q ist, daß also gilt:

$$(*) \qquad \langle Pb\rangle + \langle Qb\rangle + PbcQbc \rightarrow PcQc.$$

Die Reihe $PbcQbc$ kann man wegen $(Qb)^2$ auch als $PbcbQc$ schreiben nd hat $bcb = PcQ$ und daraus $(PcQ)^2$. Nun ist wegen $(PcQ)^2$ auch $= QPcQP$ und $c\mathfrak{o} = 1 = QPcQPQPcQP$ oder $cQPQPcQPQP$ oder $QPPPQPbbcQPQP$ oder wegen der Vertauschbarkeit von b mit P und Q uch $cQPPbPQPbcQPQP$. Letzteres kann man aber schreiben:

$$\langle cQP\rangle\langle Pb\rangle PQ\langle Pb\rangle\langle cQP\rangle QP.$$

Nach Axiom C folgt hieraus $(cQPPbP)^2$ oder $cQbPcQbP$ oder egen $PcQ = QcP$ auch $cQbQcPbP$. Nun ist aber $QbQ = b$ und $PbP = P$, also hat man $cbcb$. Man bekommt daher aus der letzten der oraussetzungen $(*)$ $PcbQbc$ oder $PcQc$, was zu beweisen war. Jetzt ann man weiter den Satz beweisen:

Für drei Punkte gilt immer $PQRPQR$ aber nie PQR, den wir üher als erstes Zusatzaxiom des § 5 formuliert, den wir aber hier im nhang nicht als Axiom gefordert haben. Es gibt nämlich zu P und R, ie wir soeben sahen, ein c mit $PcRc$ oder $P = cRc$. Ferner gibt es zu Q und R nach Axiom D ein a mit $(Qa)^2$ und $(Ra)^2$. Nach Axiom B ilt jetzt $[(\langle Ra\rangle ac)^2 Q]$ oder $[RcRcQ]$ oder $[RPQ]$, w. z. b. w.

Mit den vier Zusatzaxiomen unserer §§ 5 bis 8 gelten jetzt alle in diesen bschnitten bewiesenen Sätze. Im besonderen können wir jetzt aus Satz 52 nd Axiom D schließen, daß es zu zwei verschiedenen Punkten genau eine erbindungsgerade gibt.

Weiter kann man ziemlich leicht unter Zuhilfenahme von Sätzen der § 5 bis 8 schließen, daß man von einem Punkt auf eine Gerade stets genau n Lot fällen kann, ferner unter Zuhilfenahme von Axiom E, daß es zu einer egebenen Geraden a durch einen Punkt außerhalb stets genau eine Gerade ibt, die mit a keinen Punkt gemein hat (Parallele im gewöhnlichen Sinn) nd daß diese Gerade b zu a auch parallel im Sinn der Gruppendefinition $abZ]$ ist.

Wir können jetzt durch einen Punkt O zwei senkrechte Gerade p nd q ziehen (solche gibt es, wie leicht aus Axiom D zu sehen, weil wir

die Existenz zweier Punkte vorausgesetzt haben) und eine *Streckenrechnung* im Sinne von Hilbert aufbauen:

Wir fassen die Punkte auf p als Elemente eines Körpers auf, indem wir für sie zwei Verknüpfungen, Addition und Multiplikation definieren und zwar genau nach Maßgabe der Hilbertschen Streckenrechnung (vgl. Hilbert, Grundlagen der Geometrie, 7. Aufl., Kap. III, § 15). Als die Summe zweier Punkte P und Q auf p definieren wir den gleichfalls auf p gelegenen Punkt $[QOP]$. (Es sind ja im Sinne der gewöhnlichen Geometrie die Strecken von O nach P und von Q nach $[QOP]$ kongruent, wie die Definition der vier im Parallelogramm gelegenen Punkte für den speziellen Fall lehrt, daß das ganze Parallelogramm in eine Gerade fällt.) Die Multiplikation definieren wir wörtlich nach Hilbert. Alle dortigen Konstruktionen können wir nämlich jetzt eindeutig ausführen. Für die beiden Verknüpfungen gelten dann die sämtlichen Körperaxiome, da im besonderen auch der Pascalsche Satz für ein Sechseck mit seinen Ecken auf zwei senkrechten Geraden beweisbar ist. Dieser Satz kann ja auf den Höhensatz zurückgeführt werden.

Wir können somit zu unserer Gruppe einen Körper hinzukonstruieren, und es läßt sich dann weiter zeigen, daß sich den Punkt- und Geradenspiegelungen der analytischen Geometrie dieses Körpers (vgl. (4) und (5) der Einleitung), derart eindeutig umkehrbar die Punkte und Geraden (involutorischen Elemente) unserer Gruppe zuordnen lassen, daß sich die in der Körpergeometrie aus den Punkt- und Geradenspiegelungen erzeugte Gruppe der kongruenten Abbildungen auf die aus den involutorischen Elementen erzeugte Gruppe, von der wir ausgegangen sind, isomorph abbildet.

(Eingegangen am 24. Dezember 1930.)

Über die Ableitungsgleichungen der ametrischen Mannigfaltigkeiten.

Von

Josef Lense in München.

Die Mannigfaltigkeiten des euklidischen R_n, deren Bogenelement verschwindet, sind der Gegenstand der Untersuchung mehrerer Arbeiten des Verfassers gewesen[1]). Ich habe für sie den Namen ametrisch verwendet, von anderer Seite wird dafür total- oder vollisotrop vorgeschlagen. In der vorliegenden Arbeit sollen ihre Ableitungsgleichungen aufgestellt und genauer untersucht werden. Daraus werden Folgerungen über die Mannigfaltigkeit der Tangentialräume ametrischer Mannigfaltigkeiten gezogen. Ein Beispiel aus der Theorie dieser Mannigfaltigkeiten (die allgemeinste ametrische Fläche (V_2) des euklidischen R_n, zugleich die Integration der Differentialgleichung $\sum_{\nu=1}^{6} dx_\nu^2 = 0$) beschließt die Arbeit.

§ 1.

Das begleitende n-Bein[2]).

Wir untersuchen eine ametrische Mannigfaltigkeit V_μ von der Dimension μ in der Umgebung eines ihrer regulären Punkte P. Es muß $\mu \leqq \left[\frac{n}{2}\right]$ sein[3]). Wir setzen $m = n - 2\mu$. Der Tangential-R_μ der V_μ trifft die

[1]) J. Lense, I. Über ametrische Mannigfaltigkeiten und quadratische Differentialformen mit verschwindender Diskriminante, Deutsche Math.-Ver. **35** (1926), S. 280—294; II. Über spezielle ametrische Mannigfaltigkeiten, Wien. Ber. **135** (1926), S. 29—32; II. Ein Beitrag zur Kugelgeometrie, Wien. Ber. **136** (1927), S. 81—85; IV. Über die Tangentialräume ametrischer Mannigfaltigkeiten, Math. Zeitschr. **29** (1928), S. 87—95. — Die Arbeiten sind im folgenden mit I—IV bezeichnet.

[2]) Den Gedanken, dieses n-Bein zu wählen, und überhaupt die Anregung und verschiedene Ratschläge zu dieser Arbeit verdanke ich Herrn Weyl.

[3]) Über die Existenz derartiger Mannigfaltigkeiten vgl. I, S. 283.

absolute nicht-singuläre Mannigfaltigkeit zweiter Ordnung A^2_{n-2}, die im uneigentlichen R^u_{n-1} des R_n liegt und die euklidische Maßbestimmung des R_n bewirkt, in einem linearen $R^u_{\mu-1}$. Der obere Stellenzeiger u bezieht sich immer darauf, daß das betreffende Gebilde uneigentlich ist. Wir wählen in dem Tangential-R_μ μ linear unabhängige, von P ausgehende Vektoren $\mathfrak{x}_\alpha$ $(\alpha = 1, 2, \ldots, \mu)$. Ihre uneigentlichen Punkte bezeichnen wir mit x^u_α. Sie sind μ linear unabhängige Punkte des $R^u_{\mu-1}$ und zu sich selbst und untereinander konjugiert, also gilt

$$\mathfrak{x}_\alpha \mathfrak{x}_\beta = 0 \qquad (\alpha, \beta = 1, 2, \ldots, \mu). \tag{1}$$

Auf einem anderen $R^u_{\mu-1}$ der A^2_{n-2} (sie enthält unendliche viele) und außerhalb des eben benützten (wir bezeichnen ihn zum Unterschied mit einem Querstrich) wählen wir weitere μ linear unabhängige Punkte $\bar{x}^u_\alpha$ (sie hängen daher auch von den x^u_α nicht linear ab) und erhalten damit weitere μ von P ausgehende Vektoren $\bar{\mathfrak{x}}_\alpha$, deren uneigentliche Punkte sie sind. Analog wie früher gilt

$$\bar{\mathfrak{x}}_\alpha \bar{\mathfrak{x}}_\beta = 0. \tag{2}$$

Durch passende Wahl der $\bar{x}^u_\alpha$ können wir noch erreichen, daß x^u_α und $\bar{x}^u_\beta$ konjugiert sind, falls $\alpha \neq \beta$. Die Tangential-R^u_{n-2} der $\mu - 1$ Punkte x^u_α $(\alpha \neq \beta)$ bezüglich der A^2_{n-2} schneiden nämlich den $\bar{R}^u_{n-1}$ in einem Punkt $\bar{x}^u_\beta$, der gerade die verlangte Eigenschaft hat. Wir erhalten daher bei passender Normierung als drittes Gleichungssystem

$$\mathfrak{x}_\alpha \bar{\mathfrak{x}}_\beta = \delta_{\alpha\beta} \qquad \left(\delta_{\alpha\beta} = \begin{cases} 0, & \text{wenn } \alpha \neq \beta \\ 1, & \text{wenn } \alpha = \beta \end{cases}\right). \tag{3}$$

Damit sind 2μ linear unabhängige Vektoren im Punkte P gefunden. Wir ergänzen sie durch m weitere Vektoren $\mathfrak{n}_i$ $(i = 1, 2, \ldots, m)$ in folgender Weise zu einem n-Bein des R_n: Die Tangential-R^u_{n-2} der 2μ Punkte $x^u_\alpha, \bar{x}^u_\alpha$ schneiden sich in einem R^u_{m-1}. Auf diesem wählen wir außerhalb der A^2_{n-2} m linear unabhängige Punkte n^u_i als uneigentliche Punkte der Vektoren $\mathfrak{n}_i$. Da die n^u_i zu den $x^u_\alpha, \bar{x}^u_\alpha$ konjugiert sind, erhalten wir bei passender Normierung die drei weiteren Gleichungssysteme:

$$\mathfrak{x}_\alpha \mathfrak{n}_i = 0, \tag{4}$$

$$\bar{\mathfrak{x}}_\alpha \mathfrak{n}_i = 0, \tag{5}$$

$$\mathfrak{n}_i \mathfrak{n}_k = \delta_{ik} \qquad (i, k = 1, 2, \ldots, m). \tag{6}$$

Wie man leicht sieht, legen die Vektoren $\mathfrak{x}_\alpha$, $\mathfrak{n}_i$ den Normalraum $N_{n-\mu}$ der V_μ im Punkte P fest. Denken wir uns die V_μ durch ihre Vektorgleichung

$$\mathfrak{x} = \mathfrak{x}(u_1, u_2, \ldots, u_\mu) \tag{7}$$

gegeben, so können wir für die $\mathfrak{x}_\alpha$ die Ableitungen des Vektors $\mathfrak{x}$ nach den Parametern u_α wählen, also

$$\mathfrak{x}_\alpha = \frac{\partial \mathfrak{x}}{\partial u_\alpha}. \tag{8}$$

Gleichzeitig mit den $\mathfrak{x}_\alpha$ werden auch die $\bar{\mathfrak{x}}_\alpha$ und $\mathfrak{n}_i$ Funktionen der u_α und sollen ebenso wie $\mathfrak{x}$ regulär analytisch in der Umgebung von P vorausgesetzt werden. Differentiationen dieser Vektoren nach den Parametern u_α bezeichnen wir im folgenden mit dem Stellenzeiger α, setzen aber zwischen den Stellenzeigern, die den Vektor numerieren, und denen, die auf die Differentiation hinweisen, einen Beistrich. Es ist also zum Beispiel

$$\bar{\mathfrak{x}}_{\alpha,\beta} = \frac{\partial \bar{\mathfrak{x}}_\alpha}{\partial u_\beta}, \quad \bar{\mathfrak{x}}_{\alpha,\beta\gamma} = \frac{\partial^2 \bar{\mathfrak{x}}_\alpha}{\partial u_\beta \partial u_\gamma}, \quad \mathfrak{n}_{i,\beta} = \frac{\partial \mathfrak{n}_i}{\partial u_\beta}, \quad \mathfrak{n}_{i,\beta\gamma} = \frac{\partial^2 \mathfrak{n}_i}{\partial u_\beta \partial u_\gamma}.$$

Dagegen schreiben wir kurz

$$\frac{\partial^2 \mathfrak{x}}{\partial u_\alpha \partial u_\beta} = \mathfrak{x}_{\alpha\beta}, \qquad \frac{\partial^3 \mathfrak{x}}{\partial u_\alpha \partial u_\beta \partial u_\gamma} = \mathfrak{x}_{\alpha\beta\gamma}.$$

So ergeben sich aus den Beziehungen (1) bis (6) die Gleichungen[4])

$$\begin{aligned}
\mathfrak{x}_\alpha \mathfrak{x}_{\beta\gamma} &= 0, \\
\bar{\mathfrak{x}}_{\alpha,\gamma} \bar{\mathfrak{x}}_\beta + \bar{\mathfrak{x}}_\alpha \bar{\mathfrak{x}}_{\beta,\gamma} &= 0, \\
\mathfrak{x}_{\alpha\gamma} \bar{\mathfrak{x}}_\beta + \mathfrak{x}_\alpha \mathfrak{x}_{\beta,\gamma} &= 0, \qquad (\alpha, \beta, \gamma = 1, 2, \ldots, \mu; \\
\mathfrak{x}_{\alpha\beta} \mathfrak{n}_i + \mathfrak{x}_\alpha \mathfrak{n}_{i,\beta} &= 0, \qquad i, k = 1, 2, \ldots, m). \\
\bar{\mathfrak{x}}_{\alpha,\beta} \mathfrak{n}_i + \bar{\mathfrak{x}}_\alpha \mathfrak{n}_{i,\beta} &= 0, \\
\mathfrak{n}_{i,\beta} \mathfrak{n}_k + \mathfrak{n}_i \mathfrak{n}_{k,\beta} &= 0.
\end{aligned} \tag{9}$$

Es möge hier noch auf eine andere Charakterisierung der ametrischen Mannigfaltigkeiten hingewiesen werden, die aus den von W. Mayer[5]) eingeführten Begriffen folgt. Herr Mayer bezeichnet den von den Ableitungen des Ortsvektors $\mathfrak{x}$ bis zur h-ten Ordnung aufgespannten Schmiegraum mit $J_{12\ldots h}$, dagegen denjenigen seiner Untervektorräume, der auf $J_{12\ldots h-1}$ senkrecht steht, mit J_h und nennt J_h den h-ten Normalraum. J_2 ist also der Untervektorraum des Schmiegraumes J_{12}, der auf dem Tangentialraum J_1 senkrecht steht, J_3 der Untervektorraum von J_{123}, der auf dem Schmiegraum J_{12} senkrecht steht, usw. Da nach IV., S. 89, *eine ametrische Mannigfaltigkeit dadurch charakterisiert werden kann, daß der Schmiegraum J_{12} auf dem Tangentialraum J_1 senkrecht steht, läßt sich diese Bedingung auch durch folgende ersetzen*:

---

[4]) Bezüglich der ersten der Gleichungen (9) vgl. IV, S. 89.

[5]) A. Duschek u. W. Mayer, Lehrbuch der Differentialgeometrie 2, S. 201, Leipzig u. Berlin 1930, B. G. Teubner.

1. *J_2 fällt mit J_{12} zusammen.*

2. *J_1 liegt in J_2.*

Wie man sich leicht überzeugt, ist jede dieser beiden Bedingungen für sich notwendig und hinreichend dafür, daß die Mannigfaltigkeit ametrisch ist. Über die höheren Normalräume läßt sich im allgemeinen nichts aussagen.

§ 2.

Die Ableitungsgleichungen.

Da die Vektoren $\mathfrak{x}_\alpha$, $\bar{\mathfrak{x}}_\alpha$, $\mathfrak{n}_i$ ein Koordinatensystem des R_n im Punkte P bilden, müssen sich ihre Ableitungen nach den Parametern u_β linear homogen durch sie darstellen lassen, wir können also ansetzen

$$
\begin{aligned}
\mathfrak{x}_{\alpha\beta} &= \Gamma^\lambda_{\alpha\beta}\mathfrak{x}_\lambda + b^s_{\alpha\beta}\mathfrak{n}_s,\\
\bar{\mathfrak{x}}_{\alpha,\beta} &= \bar{\Gamma}^\lambda_{\alpha\beta}\mathfrak{x}_\lambda + \bar{b}^s_{\alpha\beta}\mathfrak{n}_s + d^\lambda_{\alpha\beta}\bar{\mathfrak{x}}_\lambda,\\
\mathfrak{n}_{i,\beta} &= a^\lambda_{i\beta}\mathfrak{x}_\lambda + c^s_{i\beta}\mathfrak{n}_s + e^\lambda_{i\beta}\bar{\mathfrak{x}}_\lambda.
\end{aligned}
\tag{10}
$$

Die Bezeichnung ist im folgenden immer so gewählt, daß über doppelt auftretende Stellenzeiger in bekannter Weise nach A. Einstein summiert wird, und zwar bei griechischen von 1 bis μ, bei lateinischen von 1 bis m.

Aus den Gleichungen (8) und (9) ergeben sich folgende Beziehungen zwischen den Koeffizienten der Vektoren:

$$
\begin{gathered}
\Gamma^\lambda_{\alpha\beta} = \Gamma^\gamma_{\beta\alpha}, \qquad \bar{\Gamma}^\beta_{\alpha\gamma} + \bar{\Gamma}^\alpha_{\beta\gamma} = 0,\\
b^\gamma_{\alpha\beta} = b^\gamma_{\beta\alpha}, \qquad b^i_{\alpha\beta} + e^\alpha_{i\beta} = 0,\\
\Gamma^\beta_{\alpha\gamma} + d^\alpha_{\beta\gamma} = 0, \qquad \bar{b}^i_{\alpha\beta} + a^\alpha_{i\beta} = 0,\\
c^k_{i\beta} + c^i_{k\beta} = 0,
\end{gathered}
\tag{11}
$$

$$
\begin{gathered}
\Gamma^\gamma_{\alpha\beta} = \bar{\mathfrak{x}}_\gamma\mathfrak{x}_{\alpha\beta}, \qquad b^i_{\alpha\beta} = \mathfrak{n}_i\mathfrak{x}_{\alpha\beta},\\
\bar{\Gamma}^\gamma_{\alpha\beta} = \mathfrak{x}_\gamma\bar{\mathfrak{x}}_{\alpha,\beta}, \qquad \bar{b}^i_{\alpha\beta} = \mathfrak{n}_i\bar{\mathfrak{x}}_{\alpha,\beta},\\
c^k_{i\alpha} = \mathfrak{n}_k\mathfrak{n}_{i,\alpha}.
\end{gathered}
$$

Wir können also statt (10) schreiben:

$$
\begin{aligned}
\mathfrak{x}_{\alpha\beta} &= \Gamma^\lambda_{\alpha\beta}\mathfrak{x}_\lambda + b^s_{\alpha\beta}\mathfrak{n}_s,\\
\bar{\mathfrak{x}}_{\alpha,\beta} &= \bar{\Gamma}^\lambda_{\alpha\beta}\mathfrak{x}_\lambda - \Gamma^\alpha_{\lambda\beta}\bar{\mathfrak{x}}_\lambda + \bar{b}^s_{\alpha\beta}\mathfrak{n}_s,\\
\mathfrak{n}_{i,\beta} &= -\bar{b}^i_{\lambda\beta}\mathfrak{x}_\lambda - b^i_{\lambda\beta}\bar{\mathfrak{x}}_\lambda + c^s_{i\beta}\mathfrak{n}_s.
\end{aligned}
\tag{12}
$$

Dabei sind die Koeffizienten regulär analytisch in der Umgebung von P

vorausgesetzt, und zwischen ihnen bestehen die Beziehungen:

$$\Gamma^{\nu}_{\alpha\beta} = \Gamma^{\nu}_{\beta\alpha}, \qquad b^{i}_{\alpha\beta} = b^{i}_{\beta\alpha},$$
$$\overline{\Gamma}^{\beta}_{\alpha\gamma} + \overline{\Gamma}^{\alpha}_{\beta\gamma} = 0, \qquad c^{k}_{i\beta} + c^{i}_{k\beta} = 0. \tag{13}$$

Außerdem aber müssen sie noch den aus den Integrabilitätsbedingungen von (12) folgenden Gleichungen genügen. Wir erhalten sie in Analogie zur Flächentheorie in folgender Weise: Wir differenzieren die Gleichungen (12) nach u_γ und drücken die Ableitungen der rechten Seiten wieder mit Hilfe von (12) durch die Vektoren unseres Koordinatensystems aus. Vertauschen wir dann β mit γ und vergleichen die Koeffizienten von $\mathfrak{x}_\alpha$, $\mathfrak{x}_{\alpha\beta}$, $\mathfrak{n}_i$, so ergeben sich folgende Gleichungen, in denen auch bei den Koeffizienten die Differentiation nach u_α wie bei den Vektoren bloß durch den Stellenzeiger α nach dem Beistrich bezeichnet ist:

$$b^{\varkappa}_{\alpha\beta}\, b^{\varkappa}_{\delta\gamma} - b^{\varkappa}_{\alpha\gamma}\, b^{\varkappa}_{\delta\beta} = 0, \tag{14}$$

$$\Gamma^{\delta}_{\alpha\beta,\gamma} - \Gamma^{\delta}_{\alpha\gamma,\beta} + \Gamma^{\lambda}_{\alpha\beta}\,\Gamma^{\delta}_{\lambda\gamma} - \Gamma^{\lambda}_{\alpha\gamma}\,\Gamma^{\delta}_{\lambda\beta} - b^{\varkappa}_{\alpha\beta}\,\overline{b}^{\varkappa}_{\delta\gamma} + b^{\varkappa}_{\alpha\gamma}\,\overline{b}^{\varkappa}_{\delta\beta} = 0, \tag{15}$$

$$b^{i}_{\alpha\beta,\gamma} - b^{i}_{\alpha\gamma,\beta} + \Gamma^{\lambda}_{\alpha\beta}\, b^{i}_{\lambda\gamma} - \Gamma^{\lambda}_{\alpha\gamma}\, b^{i}_{\lambda\beta} + b^{\varkappa}_{\alpha\beta}\, c^{i}_{\varkappa\gamma} - b^{\varkappa}_{\alpha\gamma}\, c^{i}_{\varkappa\beta} = 0, \tag{16}$$

$$\overline{\Gamma}^{\delta}_{\alpha\beta,\gamma} - \overline{\Gamma}^{\delta}_{\alpha\gamma,\beta} + \overline{\Gamma}^{\lambda}_{\alpha\beta}\,\Gamma^{\delta}_{\lambda\gamma} - \overline{\Gamma}^{\lambda}_{\alpha\gamma}\,\Gamma^{\delta}_{\lambda\beta} - \Gamma^{\alpha}_{\lambda\beta}\,\overline{\Gamma}^{\delta}_{\lambda\gamma} + \Gamma^{\alpha}_{\lambda\gamma}\,\overline{\Gamma}^{\delta}_{\lambda\beta} - \overline{b}^{\varkappa}_{\alpha\beta}\,\overline{b}^{\varkappa}_{\delta\gamma} + \overline{b}^{\varkappa}_{\alpha\gamma}\,\overline{b}^{\varkappa}_{\delta\beta} = 0, \tag{17}$$

$$\overline{b}^{i}_{\alpha\beta,\gamma} - \overline{b}^{i}_{\alpha\gamma,\beta} + \overline{\Gamma}^{\lambda}_{\alpha\beta}\, b^{i}_{\lambda\gamma} - \overline{\Gamma}^{\lambda}_{\alpha\gamma}\, b^{i}_{\lambda\beta} - \Gamma^{\alpha}_{\lambda\beta}\,\overline{b}^{i}_{\lambda\gamma} + \Gamma^{\alpha}_{\lambda\gamma}\,\overline{b}^{i}_{\lambda\beta} + \overline{b}^{\varkappa}_{\alpha\beta}\, c^{i}_{\varkappa\gamma} - \overline{b}^{\varkappa}_{\alpha\gamma}\, c^{i}_{\varkappa\beta} = 0, \tag{18}$$

$$c^{k}_{i\beta,\gamma} - c^{k}_{i\gamma,\beta} - \overline{b}^{i}_{\lambda\beta}\, b^{k}_{\lambda\gamma} + \overline{b}^{i}_{\lambda\gamma}\, b^{k}_{\lambda\beta} + c^{\varkappa}_{i\beta}\, c^{k}_{\varkappa\gamma} - c^{\varkappa}_{i\gamma}\, c^{k}_{\varkappa\beta} - b^{i}_{\lambda\beta}\,\overline{b}^{k}_{\lambda\gamma} + b^{i}_{\lambda\gamma}\,\overline{b}^{k}_{\lambda\beta} = 0. \tag{19}$$

Zu den zwischen den Koeffizienten bestehenden algebraischen Beziehungen (13) kommen also noch die Gleichungen (14) als rein algebraische Gleichungen hinzu, die übrigen (15) bis (19) sind die Differentialgleichungen, denen die Koeffizienten genügen müssen. Sie sind das Analogon der Gauß-schen und Codazzi-Mainardischen Gleichungen der Flächentheorie. Wählt man die Koeffizienten so als regulär analytische Funktionen der Parameter u_α in der Umgebung einer bestimmten Stelle $(u_1^0, u_2^0, \ldots, u_\mu^0)$, daß sie den Beziehungen (13) bis (19) genügen, so erweisen sich die Differentialgleichungen (8) und (12) integrabel und liefern $\mathfrak{x}$, $\mathfrak{x}_\alpha$, $\mathfrak{n}_i$ als regulär analytische Funktionen in der Umgebung dieser Stelle. Dabei bleiben nach bekannten Sätzen aus der Theorie der Differentialgleichungen dort die Werte von $\mathfrak{x}$, $\mathfrak{x}_\alpha$, $\mathfrak{n}_i$ willkürlich. Wählt man diese Anfangswerte so, daß sie

den Beziehungen (1) bis (6) genügen, so zeigt sich, daß diese Beziehunge in allen Punkten einer passenden Umgebung der Stelle erfüllt sind. Den mit Hilfe der Ableitungsgleichungen (12) folgt:

$$
\begin{aligned}
\frac{\partial(\mathfrak{x}_\alpha \mathfrak{x}_\beta)}{\partial u_\gamma} &= \Gamma^\lambda_{\alpha\gamma}\,\mathfrak{x}_\lambda \mathfrak{x}_\beta + \Gamma^\lambda_{\beta\gamma}\,\mathfrak{x}_\lambda \mathfrak{x}_\alpha + b^s_{\alpha\gamma}\,\mathfrak{n}_s \mathfrak{x}_\beta + b^s_{\beta\gamma}\,\mathfrak{n}_s \mathfrak{x}_\alpha,\\
\frac{\partial(\bar{\mathfrak{x}}_\alpha \bar{\mathfrak{x}}_\beta)}{\partial u_\gamma} &= \bar{\Gamma}^\lambda_{\alpha\gamma}\,\mathfrak{x}_\lambda \bar{\mathfrak{x}}_\beta + \bar{\Gamma}^\lambda_{\beta\gamma}\,\mathfrak{x}_\lambda \mathfrak{x}_\alpha - \Gamma^\alpha_{\lambda\gamma}\,\bar{\mathfrak{x}}_\lambda \bar{\mathfrak{x}}_\beta - \Gamma^\beta_{\lambda\gamma}\,\bar{\mathfrak{x}}_\lambda \bar{\mathfrak{x}}_\alpha\\
&\qquad + \bar{b}^s_{\alpha\gamma}\,\mathfrak{n}_s \bar{\mathfrak{x}}_\beta + \bar{b}^s_{\beta\gamma}\,\mathfrak{n}_s \bar{\mathfrak{x}}_\alpha,\\
\frac{\partial(\mathfrak{x}_\alpha \bar{\mathfrak{x}}_\beta)}{\partial u_\gamma} &= \bar{\Gamma}^\lambda_{\beta\gamma}\,\mathfrak{x}_\lambda \mathfrak{x}_\alpha + \Gamma^\lambda_{\alpha\gamma}\,\mathfrak{x}_\lambda \bar{\mathfrak{x}}_\beta - \Gamma^\beta_{\lambda\gamma}\,\bar{\mathfrak{x}}_\lambda \mathfrak{x}_\alpha + \bar{b}^s_{\beta\gamma}\,\mathfrak{n}_s \mathfrak{x}_\alpha + b^s_{\alpha\gamma}\,\mathfrak{n}_s \bar{\mathfrak{x}}_\beta,\\
\frac{\partial(\mathfrak{x}_\alpha \mathfrak{n}_i)}{\partial u_\beta} &= -\bar{b}^i_{\lambda\beta}\,\mathfrak{x}_\lambda \mathfrak{x}_\alpha - b^i_{\lambda\beta}\,\bar{\mathfrak{x}}_\lambda \mathfrak{x}_\alpha + \Gamma^\lambda_{\alpha\beta}\,\mathfrak{x}_\lambda \mathfrak{n}_i + c^s_{i\beta}\,\mathfrak{n}_s \mathfrak{x}_\alpha + b^s_{\alpha\beta}\,\mathfrak{n}_s \mathfrak{n}_i,\\
\frac{\partial(\bar{\mathfrak{x}}_\alpha \mathfrak{n}_i)}{\partial u_\beta} &= -\bar{b}^i_{\lambda\beta}\,\mathfrak{x}_\lambda \bar{\mathfrak{x}}_\alpha - b^i_{\lambda\beta}\,\bar{\mathfrak{x}}_\lambda \bar{\mathfrak{x}}_\alpha + \bar{\Gamma}^\lambda_{\alpha\beta}\,\mathfrak{x}_\lambda \mathfrak{n}_i - \Gamma^\alpha_{\lambda\beta}\,\bar{\mathfrak{x}}_\lambda \mathfrak{n}_i\\
&\qquad + c^s_{i\beta}\,\mathfrak{n}_s \bar{\mathfrak{x}}_\alpha + \bar{b}^s_{\alpha\beta}\,\mathfrak{n}_s \mathfrak{n}_i,\\
\frac{\partial(\mathfrak{n}_i \mathfrak{n}_k)}{\partial u_\alpha} &= -b^i_{\lambda\alpha}\,\mathfrak{x}_\lambda \mathfrak{n}_k - b^k_{\lambda\alpha}\,\mathfrak{x}_\lambda \mathfrak{n}_i - b^i_{\lambda\alpha}\,\bar{\mathfrak{x}}_\lambda \mathfrak{n}_k - b^k_{\lambda\alpha}\,\bar{\mathfrak{x}}_\lambda \mathfrak{n}_i\\
&\qquad + c^s_{i\alpha}\,\mathfrak{n}_s \mathfrak{n}_k + c^s_{k\alpha}\,\mathfrak{n}_s \mathfrak{n}_i.
\end{aligned}
\tag{20}
$$

Dieses System von Differentialgleichungen für die Produkte $\mathfrak{x}_\alpha \mathfrak{x}_\beta$, $\bar{\mathfrak{x}}_\alpha \mathfrak{x}_\beta$, $\mathfrak{x}_\alpha \bar{\mathfrak{x}}_\beta$, $\mathfrak{x}_\alpha \mathfrak{n}_i$, $\bar{\mathfrak{x}}_\alpha \mathfrak{n}_i$, $\mathfrak{n}_i \mathfrak{n}_k$ hat aber gerade die Lösung (1) bis (6) und daher nach dem Eindeutigkeitssatz aus der Theorie der Differentialgleichungen nur diese, wenn man die Anfangswerte entsprechend (1) bis (6) wählt. *Eine ametrische Mannigfaltigkeit ist also durch die Koeffizienten ihrer Ableitungsgleichungen, falls sie nur den Beziehungen* (13) *bis* (19) *genügen bis auf ihre „Lage" im R_n vollständig bestimmt, d. h. bis auf Kongruenztransformationen und Spiegelungen.* Dieser Satz ist das Analogon des bekannten Bonnetschen Satzes der Flächentheorie[6]).

§ 3.

Die Integrabilitätsbedingungen.

Die Gleichungen (13) bis (19) sind sicher immer erfüllbar, da es ja ametrische Mannigfaltigkeiten gibt (vgl. Anmerkung [3])). Wir wollen jetzt das System dieser Gleichungen näher betrachten. Die Gleichungen (15) sind die Differentialgleichungen für die $\Gamma^\lambda_{\alpha\beta}$, (16) für die $b^i_{\alpha\beta}$, (17) für die $\bar{\Gamma}^\gamma_{\alpha\beta}$, (18) für die $b^i_{\alpha\beta}$, (19) für die $c^k_{i\alpha}$. Zwischen diesen Größen müssen noch die algebraischen Gleichungen (13) und (14) bestehen, wovon die ersteren gewisse Symmetriebeziehungen aussagen.

[6]) A. Duschek u. W. Mayer, Lehrbuch der Differentialgeometrie **1**, S. 178; **2**, S. 218

Wir betrachten zuerst die Gleichungen (15). Sie lassen sich, wenn man die Ableitungen der $\Gamma^{\gamma}_{\alpha\beta}$ auf die linke, alle übrigen Glieder auf die rechte Seite bringt, in folgender Form schreiben:

21) $$\Gamma^{\delta}_{\alpha\beta,\gamma} - \Gamma^{\delta}_{\alpha\gamma,\beta} = F^{\delta}_{\alpha\beta\gamma}.$$

$F^{\delta}_{\alpha\beta\gamma}$ enthält keine Differentialquotienten mehr und genügt, wie man durch eine längere Rechnung feststellt, folgenden Beziehungen:

22) $$\begin{aligned} F^{\delta}_{\alpha\beta\gamma} + F^{\delta}_{\alpha\gamma\beta} &= 0, \\ F^{\delta}_{\alpha\beta\gamma} + F^{\delta}_{\beta\gamma\alpha} + F^{\delta}_{\gamma\alpha\beta} &= 0, \\ F^{\delta}_{\alpha\beta\gamma,\varepsilon} + F^{\delta}_{\alpha\gamma\varepsilon,\beta} + F^{\delta}_{\alpha\varepsilon\beta,\gamma} &= 0. \end{aligned}$$

Die erste dieser Beziehungen leuchtet unmittelbar ein. Die zweite sagt aus, daß die algebraischen Beziehungen

23) $$\Gamma^{\gamma}_{\alpha\beta} = \Gamma^{\gamma}_{\beta\alpha},$$

die einzigen, denen die $\Gamma^{\gamma}_{\alpha\beta}$ genügen, mit (21) verträglich sind; es ergibt sich sofort, wenn man in (21) α mit β vertauscht. Die Beziehungen in der dritten Zeile von (22) sind die Integrabilitätsbedingungen von (21). Man erhält sie, wenn man (21) nach u_{ε} differenziert und hierauf γ mit ε vertauscht. (21) ist also integrabel.

Ganz analoge Schlüsse und Ergebnisse gelten für die Größen $b^{i}_{\alpha\beta}$. An Stelle der Gleichungen (15) und (23) treten die Gleichungen (16) und

24) $$b^{i}_{\alpha\beta} = b^{i}_{\beta\alpha}.$$

dazu kommen noch bei den $b^{i}_{\alpha\beta}$ die Gleichungen (14). Setzen wir

25) $$b^{s}_{\alpha\beta} b^{s}_{\delta\gamma} - b^{s}_{\alpha\gamma} b^{s}_{\delta\beta} = F_{\alpha\beta\gamma\delta},$$

so folgt aus (13) und (15)

26) $F_{\alpha\beta\gamma\delta,\varepsilon} + F_{\alpha\gamma\varepsilon\delta,\beta} + F_{\alpha\varepsilon\beta\delta,\gamma}$ = homogene lineare Funktion der $F_{\alpha\beta\gamma\delta}$.

Für die Größen $\overline{\Gamma}^{\gamma}_{\alpha\beta}$ gelten ganz analoge Gleichungen wie für die $\Gamma^{\gamma}_{\alpha\beta}$, nur tritt an Stelle der zweiten Zeile von (22) ein Gleichungssystem von der Gestalt

27) $$F^{\delta}_{\alpha\beta\gamma} + F^{\alpha}_{\delta\beta\gamma} = 0,$$

an Stelle von (23)

28) $$\overline{\Gamma}^{\gamma}_{\alpha\beta} + \overline{\Gamma}^{\alpha}_{\gamma\beta} = 0,$$

an Stelle von (15) tritt (17).

Analoge Schlüsse und Ergebnisse wie für die $\bar{\Gamma}^{\gamma}_{\alpha\beta}$ finden bei de Größen $c^{k}_{i\alpha}$ statt. An Stelle von (23) tritt

$$c^{k}_{i\alpha} + c^{i}_{k\alpha} = 0, \tag{29}$$

(19) an Stelle von (17).

Die Größen $\bar{b}^{i}_{\alpha\beta}$ genügen keinen algebraischen Beziehungen, sonder nur den Differentialgleichungen (18). Sie haben dieselbe Gestalt wie (21) genügen den Beziehungen in der ersten und dritten Zeile von (22) un sind daher verträglich.

Es möge noch erwähnt werden, daß man die Gleichungen (14) auch in folgender Weise erhalten kann: Differentiation der ersten der Gleichunge (9) nach u_δ und Vertauschung von γ und δ liefert

$$\mathfrak{x}_{\alpha\delta}\mathfrak{x}_{\beta\gamma} - \mathfrak{x}_{\alpha\gamma}\mathfrak{x}_{\beta\delta} = 0.$$

Daraus ergibt sich sofort (14) mit Hilfe von (1), (4), (6) und (12).

§ 4.

Die Mannigfaltigkeit der Tangentialräume.

Der Tangentialraum im Punkte P der ametrischen Mannigfaltigkei wird durch die Vektoren $\mathfrak{x}_\alpha$ aufgespannt. In der Fortschreitungsrichtun $(du_1, du_2, \ldots, du_\mu)$ ergibt sich für die Änderungen dieser Vektoren aus (12

$$d\mathfrak{x}_\alpha = \bar{\Gamma}^{\lambda}_{\alpha\beta}\mathfrak{x}_\lambda du_\beta + b^{s}_{\alpha\beta}\mathfrak{n}_s du_\beta.$$

Soll sich der Tangentialraum in dieser Fortschreitungsrichtung nicht änder so müssen die $m\,\mu$ Gleichungen

$$b^{i}_{\alpha\beta}\,du_\beta = 0 \qquad (i = 1, 2, \ldots, m;\ \alpha = 1, 2, \ldots, \mu \tag{30}$$

bestehen. Der Rang der Matrix der Koeffizienten $b^{i}_{\alpha\beta}$ sei r im Punkte P und in der betrachteten Umgebung, also $0 \leqq r \leqq \mu$. Dann hat dort da System (30) $\mu - r$ linear unabhängige Lösungen, daher die V_μ nur ∞ Tangentialräume. *Die Dimensionszahl der Mannigfaltigkeit der Tangentialräume der V_μ ist also gleich dem Rang des Systems* (30).

Es gilt folgender Satz:

Ist die Matrix $\mathfrak{M}$

$$\| \xi^{\varrho} b^{s}_{\varrho\beta} \| \qquad \begin{matrix} s = 1, 2, \ldots, m, & \text{über } \varrho \text{ wird von} \\ \beta = 1, 2, \ldots, \mu; & 1 \text{ bis } \mu \text{ summiert} \end{matrix}$$

(Bed. A) *für irgendein bestimmtes System* $(\xi^1, \xi^2, \ldots, \xi^\mu)$ *vom Ran m und $2 \leqq m < \mu$, so hat auch die durch die Spalte $b^{i}_{\alpha\beta}$ (i, α fest) e weiterte Matrix denselben Rang m. Ist der Rang von $\mathfrak{M}$ kleiner als m so kann man für den Rang der erweiterten Matrix nur behaupten, da er die Zahl m nicht übersteigt.* Der Grund dieses Satzes liegt in de Gültigkeit der Beziehung (14).

Beweis. Eine $(m+1)$-reihige Determinante der erweiterten Matrix at die Gestalt

31)
$$\begin{vmatrix} \xi^\varrho b^1_{\varrho\beta_1} & \xi^\varrho b^2_{\varrho\beta_1} & \dots & \xi^\varrho b^m_{\varrho\beta_1} & b^i_{\alpha\beta_1} \\ \xi^\varrho b^1_{\varrho\beta_2} & \xi^\varrho b^2_{\varrho\beta_2} & \dots & \xi^\varrho b^m_{\varrho\beta_2} & b^i_{\alpha\beta_2} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \xi^\varrho b^1_{\varrho\beta_{m+1}} & \xi^\varrho b^2_{\varrho\beta_{m+1}} & \dots & \xi^\varrho b^m_{\varrho\beta_{m+1}} & b^i_{\alpha\beta_{m+1}} \end{vmatrix}.$$

Wir multiplizieren die letzte Spalte mit ξ^α und subtrahieren sie von der -ten. Dadurch entfernen wir dort die Glieder mit $\varrho = \alpha$ aus den Summen. Nach dem Additionssatz der Determinanten wird daher die Determinante 31) zerlegt in eine Summe von Determinanten der Gestalt

32)
$$\xi^{\varrho_i} \begin{vmatrix} \xi^\varrho b^1_{\varrho\beta_1} & \dots & b^i_{\varrho_i\beta_1} & \dots & \xi^\varrho b^m_{\varrho\beta_1} & b^i_{\alpha\beta_1} \\ \xi^\varrho b^1_{\varrho\beta_2} & \dots & b^i_{\varrho_i\beta_2} & \dots & \xi^\varrho b^m_{\varrho\beta_2} & b^i_{\alpha\beta_2} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \xi^\varrho b^1_{\varrho\beta_{m+1}} & \dots & b^i_{\varrho_i\beta_{m+1}} & \dots & \xi^\varrho b^m_{\varrho\beta_{m+1}} & b^i_{\alpha\beta_{m+1}} \end{vmatrix}.$$

In dieser Determinante ist über ϱ_i nicht zu summieren, sondern ϱ_i bedeutet einen festen von α verschiedenen Stellenzeiger. Wir entwickeln die Determinante (32) nach den Unterdeterminanten der i-ten und letzten Spalte. Sie haben die Gestalt:

33)
$$b^i_{\varrho_i\beta_\sigma} b^i_{\alpha\beta_\tau} - b^i_{\varrho_i\beta_\tau} b^i_{\alpha\beta_\sigma} \quad \text{(nicht über } i \text{ summieren)}.$$

An Stelle von (33) kann man nach (14) setzen

34)
$$-\sum_{s=1}^{m} (b^s_{\varrho_i\beta_\sigma} b^s_{\alpha\beta_\tau} - b^s_{\varrho_i\beta_\tau} b^s_{\alpha\beta_\sigma}) \qquad (s \neq i),$$

daher an Stelle der Determinante in (32) eine Summe von Determinanten derselben Gestalt, wo jedoch in der σ-ten Zeile der i-ten Spalte $b^s_{\varrho_i\beta_\sigma}$ und in der σ-ten Zeile der letzten Spalte $b^s_{\alpha\beta_\sigma}$ $(\sigma = 1, 2, \dots, m+1)$ steht. Für 31) ergibt sich daher, wenn wir die Glieder mit $\varrho = \alpha$ in der i-ten Spalte dadurch wieder einfügen, daß wir die letzte Spalte mit ξ^α multiplizieren und zur i-ten addieren,

$$-\sum_{\substack{s=1 \\ (s \neq i)}}^{m} \begin{vmatrix} \xi^\varrho b^1_{\varrho\beta_1} & \dots & \xi^\varrho b^s_{\varrho\beta_1} & \dots & \xi^\varrho b^m_{\varrho\beta_1} & b^s_{\alpha\beta_1} \\ \xi^\varrho b^1_{\varrho\beta_2} & \dots & \xi^\varrho b^s_{\varrho\beta_2} & \dots & \xi^\beta b^m_{\varrho\beta_2} & b^s_{\alpha\beta_2} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \xi^\varrho b^1_{\varrho\beta_{m+1}} & \dots & \xi^\varrho b^s_{\varrho\beta_{m+1}} & \dots & \xi^\varrho b^m_{\varrho\beta_{m+1}} & b^s_{\alpha\beta_{m+1}} \end{vmatrix} = 0,$$

weil s die Werte $1, 2, \ldots, m$ (i ausgenommen) annimmt und daher immer zwei Spalten einander gleich werden. Alle $(m+1)$-reihigen Determinanten der erweiterten Matrix verschwinden, ihr Rang ist daher $\leqq m$.

Wir betrachten zwei Spezialfälle, die für das folgende wichtig sind:

1. Der Rang der Matrix $\| b^s_{\alpha\beta} \|$ $(\alpha, \beta = 1, 2, \ldots, \mu)$ sei 1 für jedes s, also $b^s_{\alpha\beta} = b^s_\alpha \cdot b^s_\beta$ (nicht über s summieren). In diesem Falle ist

$$\| \xi^\varrho b^s_{\varrho\beta} \| = \| \xi^\varrho b^s_\varrho \cdot b^s_\beta \|,$$

der Rang von $\mathfrak{M}$ ist daher gleich dem der Matrix $\| b^s_\beta \|$, wenn

$$\prod_{i=1}^{m} \xi^\varrho b^i_\varrho \neq 0.$$

2. $b^1_{\alpha\beta} = b_{\alpha\beta} = b_{\beta\alpha}$, $b^2_{\alpha\beta} = b_{\alpha\beta}\sqrt{-1}$, $b^i_{\alpha\beta} = 0$ für $2 \leqq i \leqq m$. In diesem Fall ist der Rang von $\mathfrak{M}$, wie man leicht sieht, $\leqq 1$. Aus 1. folgern wir sofort, indem wir $\| b^s_\beta \|$ vom Rang m und das Produkt von 0 verschieden wählen: Die Bedingung A ist mit (14) verträglich.

Wir untersuchen jetzt den Rang des Systems (30). Wir multiplizieren die α-te Gleichung mit ξ^α und summieren über α. Dadurch erhalten wir das System der m Gleichungen

$$\xi^\alpha b^i_{\alpha\beta} du_\beta = 0. \tag{35}$$

Seine Koeffizientenmatrix ist gerade die Matrix $\mathfrak{M}$.

Lassen sich Multiplikatoren $\lambda^i_{\alpha s}$ derart bestimmen, daß

$$b^i_{\alpha\beta} = \lambda^i_{\alpha s} \xi^\varrho b^s_{\varrho\beta} \qquad (2 \leqq m < \mu), \tag{36}$$

dann sind die beiden Systeme (30) und (35) gleichwertig und haben daher denselben Rang. Hat $\mathfrak{M}$ den Rang m, so ist nach dem eben bewiesenen Satze das System (36) nach den $\lambda^i_{\alpha s}$ auflösbar, daher hat (30) den Rang m. Verschwindet eine m-reihige Determinante von $\mathfrak{M}$ für ein bestimmtes System der ξ^α im Punkte P nicht, so verschwindet sie für passend gewählte ξ^α auch in einer gewissen Umgebung der Stelle nicht, d. h. $\mathfrak{M}$ und damit auch (30) hat dort den Rang m, die V_μ also ∞^m Tangentialräume. Die V_μ kann somit nur dann mehr Tangentialräume haben, wenn bei allen beliebigen ξ^α der Rang von $\mathfrak{M}$ für alle Punkte der Umgebung von P kleiner als m ist. Das ist aber eine neue Bedingung B für die $b^i_{\alpha\beta}$, weil die Bedingung A mit (14) verträglich ist. Befriedigen also die $b^i_{\alpha\beta}$ die Bedingungen (14) und (24), ohne daß sie der Bedingung B genügen, so ist immer die Bedingung A in P und in einer passenden Umgebung erfüllt, d. h. in diesem Fall hat jede ametrische V_μ für $2 \leqq m < \mu$ nur ∞^m Tangentialräume.

Ist die Bedingung B erfüllt, so kann der Rang r von (30) alle Werte wischen 0 und μ annehmen [7]). Man wählt einfach in Beispiel 2 r als en Rang von $b_{\alpha\beta}$. Das Gleichungssystem (30) reduziert sich in diesem all auf

$$b_{\alpha\beta}\, du_\beta = 0.$$

Dasselbe Beispiel zeigt auch, daß die Bedingung A wohl hinreichend, aber icht notwendig für $r = m$ ist.

In den Fällen $m = 0$ und $m = 1$ ist immer $r = m$. Denn für $m = 0$ ällt das System (30) weg, die V_μ hat überhaupt nur einen festen Tanentialraum, d. h. sie fällt mit ihm zusammen, ist also linear [8]). Für $= 1$ ist gemäß (14) $r = 1$, [9]) die V_u kann als Hüllgebilde der Schmiegäume einer Kurve des R_n aufgefaßt werden [10]).

Ist $m \geqq \mu$, so ist im allgemeinen $r = \mu$. Soll nämlich $r < \mu$ sein, müssen sämtliche μ-reihigen Determinanten der Koeffizienten des ystems (30) verschwinden (Bedingung C). Die Bedingung C hängt von 4) und (24) nicht ab, wie Beispiel 2 zeigt, wenn man die Matrix $\| b_{\alpha\beta} \|$ om Range μ wählt. Der Fall $\mu = m = 1$ liefert die ametrischen oder otropen Kurven des R_3. Sie sind entweder nicht linear und haben dann ∞^1 angenten, oder sie sind Gerade und fallen mit ihren Tangenten zusammen. ∞^μ Tangentialräume können also nur für $1 \leqq \mu < \frac{n-1}{2}$ auftreten, falls > 3 ist.

Mit den Untersuchungen des § 4 ist keineswegs die Existenz der den nzelnen Fällen entsprechenden ametrischen Mannigfaltigkeiten nachgeiesen, sondern es sind nur gewisse Fälle ausgeschieden. Zum Nachweis er Existenz müßte erst gezeigt werden, daß die in den einzelnen Fällen ber die $b^i_{\alpha\beta}$ gemachten Voraussetzungen jedesmal mit dem System (13) s (19) verträglich sind. Das wäre z. B. der Fall, wenn man dartun önnte, daß die Anfangswerte der $b^i_{\alpha\beta}$, abgesehen von den Bedingungen 4) und (24), beliebig gewählt werden können. Nimmt man die Ergebsse von I hinzu, so kann man für die Umgebung einer regulären Stelle sammenfassend folgendes aussagen: *In einem euklidischen R_n gibt es*

---

[7]) Diesen Fall habe ich in IV übersehen. Wie mich Herr Weyl aufmerksam achte, hat der dort angegebene Gedankengang keine Beweiskraft, weil sich r bei r Projektion ändern kann; vgl. IV, S. 89 u. S. 92; die Entwicklungen des § 3 der beit IV bleiben in Geltung.

[8]) Vgl. hierzu I, S. 283; II, S. 30; III, S. 82.

[9]) Vgl. hierzu I, S. 285; II, S. 30; IV, S. 92.

[10]) C. Segre, Preliminari di una teoria delle varietà luoghi di spazi, Palermo nd. **30** (1910), S. 87—121.

ametrische Mannigfaltigkeiten V_μ für $1 \leqq \mu \leqq \left[\frac{n}{2}\right]$. Hat μ seinen größten Wert $\left[\frac{n}{2}\right]$, so ist die V_μ für gerade n linear, für ungerade linear oder eine Torse mit ∞^1 Tangentialräumen.

§ 5.

Die ametrischen Flächen des R_6.

Die analytische Darstellung sämtlicher Mannigfaltigkeiten des euklidischen R_n ist identisch mit der Integration der totalen Differentialgleichung

$$\sum_{r=1}^{n} dx_r^2 = 0, \tag{37}$$

wo die x_r rechtwinkelige kartesische Koordinaten des R_n bedeuten. Wenn wir das in § 1 in einem festen Punkte P einer ametrischen V_μ definierte n-Bein als Koordinaten-n-Bein einführen, hat ein beliebiger Vektor $\mathfrak{y}$ des R_n die Gestalt

$$\mathfrak{y} = \sum_{r=1}^{\mu} (y_r \mathfrak{x}_r + \bar{y}_r \bar{\mathfrak{x}}_r) + \sum_{s=1}^{m} z_s \mathfrak{n}_s, \tag{38}$$

daher lautet das quadrierte Bogenelement

$$d\mathfrak{y}^2 = \sum_{r=1}^{\mu} dy_r\, d\bar{y}_r + \sum_{s=1}^{m} dz_s^2. \tag{39}$$

Führen wir die y_r in der Umgebung von P als Parameter ein, so erhalten wir die Gleichungen der V_μ in der Gestalt

$$\begin{aligned} \bar{y}_r &= \bar{y}_r(y_1, y_2, \ldots, y_\mu) \qquad (r = 1, 2, \ldots, \mu), \\ z_s &= z_s(y_1, y_2, \ldots, y_\mu) \qquad (s = 1, 2, \ldots, m). \end{aligned} \tag{40}$$

Die ametrischen V_μ niedrigster Dimension ($\mu > 1$) mit ∞^μ Tangentialräumen im R_n niedrigster Dimension sind nach dem § 4 die ametrischen Flächen (V_2) im R_6. Wir weisen ihre Existenz durch direkte Integration von (37) nach, ohne die Ableitungsgleichungen heranzuziehen. Nach dem eben Gesagten können wir (37) in der Form (39), die Gleichungen der Fläche in der Gestalt (40) mit $\mu = 2$, $m = 2$ voraussetzen. Führen wir noch

$$y_3 = z_1 + i z_2, \qquad \bar{y}_3 = z_1 - i z_2 \qquad (i = \sqrt{-1}) \tag{41}$$

als neue Veränderliche ein, so lautet die zu integrierende Differentialgleichung

$$dy_1\, d\bar{y}_1 + dy_2\, d\bar{y}_2 + dy_3\, d\bar{y}_3 = 0, \tag{42}$$

die Gleichungen der Fläche

$$\begin{aligned} \bar{y}_1 &= \bar{y}_1(y_1, y_3), \quad \bar{y}_2 = \bar{y}_2(y_1, y_2), \\ y_3 &= y_3(y_1, y_2), \quad \bar{y}_3 = \bar{y}_3(y_1, y_2). \end{aligned} \tag{43}$$

Auf die Form (42) und den Ansatz (43) kann man auch unabhängig on den vorangegangenen Entwicklungen in folgender Weise kommen: ührt man in (37) die neuen Veränderlichen

$$\begin{aligned} y_1 &= x_1 + i x_2, \quad \bar{y}_1 = x_1 - i x_2, \\ y_2 &= x_3 + i x_4, \quad \bar{y}_2 = x_3 - i x_4, \\ y_3 &= x_5 + i x_6, \quad \bar{y}_3 = x_5 - i x_6 \end{aligned} \qquad (i = \sqrt{-1}) \tag{44}$$

in, so ergibt sich unmittelbar die Gleichung (42). Soll man (37) durch ine Fläche integrieren, so hat man anzusetzen:

$$x_\nu = x_\nu(u, v) \qquad (\nu = 1, 2, 3, 4, 5, 6), \tag{45}$$

o die Funktionalmatrix der x_ν nach u und v in einem bestimmten Beeich vom Rang 2 (sonst wäre die Integralmannigfaltigkeit keine Fläche, ondern eine Kurve oder ein Punkt) und die x_ν dort regulär analytisch nd. Da die x_ν in (37) sämtlich gleichberechtigt sind, können wir nnehmen

$$\frac{\partial(x_1, x_2)}{\partial(u, v)} \neq 0$$

dem betrachteten Bereich, daher die beiden ersten Gleichungen nach u nd v auflösen und das Ergebnis in die übrigen einsetzen, wodurch wir erhalten

$$x_\nu = x_\nu(x_1, x_2) \qquad (\nu = 3, 4, 5, 6), \tag{46}$$

. h. wir dürfen in unserem Bereich x_1 und x_2 als Parameter auf der läche voraussetzen. In den Veränderlichen (44) schreiben sich die Gleihungen (46) derart, daß $y_2, \bar{y}_2, y_3, \bar{y}_3$ als regulär analytische Funktionen on y_1 und $\bar{y}_1$ auftreten.

Nun können zwei Fälle vorkommen: Entweder $\bar{y}_1$ kommt in keinem ieser y vor oder mindestens in einem. Der erste Fall liefert keine Lösung er Differentialgleichung (42), weil sich das Glied $dy_1\, d\bar{y}_1$ nicht weghebt. leibt also nur der zweite Fall. Wegen des symmetrischen Baues von (42) önnen wir uns die Bezeichnung immer so gewählt denken, daß y_2 sicher on $\bar{y}_1$ abhängt, also

$$y_2 = y_2(y_1, \bar{y}_1) \tag{47}$$

oraussetzen. Da y_2 von $\bar{y}_1$ wirklich abhängt, gibt es in unserem Bereich nen Teilbereich, wo $\frac{\partial y_2}{\partial \bar{y}_1} \neq 0$ ist. Wir können dort (47) nach $\bar{y}_1$ auflösen nd das Ergebnis in $\bar{y}_2, y_3, \bar{y}_3$ einsetzen. Dadurch erhalten wir den Antz (43).

Wir können somit ansetzen:

$$\begin{aligned} dy_3 &= A\,dy_1 + B\,dy_2, \quad d\bar{y}_3 = C\,dy_1 + D\,dy_2, \\ d\bar{y}_1 &= E\,dy_1 + F\,dy_2, \quad d\bar{y}_2 = G\,dy_1 + H\,dy_2, \end{aligned} \tag{48}$$

wo die Koeffizienten von dy_1 und dy_2 regulär analytische Funktionen von y_1 und y_2 sind. Einsetzen in die Differentialgleichung (42) liefert

$$(49) \qquad G = -AC, \quad H = -(E + AD + BC), \quad F = -BD,$$

weil die Koeffizienten von dy_1^2, $dy_1 dy_2$, dy_2^2 verschwinden müssen.

Die Integrabilitätsbedingungen von (48) lauten, wenn die partiellen Differentialquotienten nach y_1 und y_2 durch entsprechende Stellenzeiger bezeichnet werden,

$$(50) \qquad A_{y_2} = B_{y_1}, \quad C_{y_2} = D_{y_1}, \quad E_{y_2} = F_{y_1}, \quad G_{y_2} = H_{y_1}.$$

Aus (49) und (50) folgt

$$(51) \qquad E_{y_1} = -A_{y_1} D - BC_{y_1}, \quad E_{y_2} = -A_{y_2} D - BC_{y_2},$$

daraus durch Gleichsetzen der gemischten Ableitungen

$$A_{y_1} D_{y_2} + B_{y_2} C_{y_1} = 2 A_{y_2} C_{y_2}.$$

Damit sind auch E_{y_1} und E_{y_2} eliminiert, und es bleibt, sobald $B_{y_2} \neq 0$ ist, das System

$$(52) \qquad B_{y_1} = A_{y_2}, \quad C_{y_1} = \frac{2 A_{y_2} C_{y_2} - A_{y_1} D_{y_2}}{B_{y_2}}, \quad D_{y_1} = C_{y_2}.$$

Es ist aufgelöst nach den partiellen Ableitungen der drei unbekannten Funktionen B, C, D nach derselben Veränderlichen y_1 und daher nach dem Satz von A. Cauchy und S. Kowalewski[11]) verträglich. Die Integration ergibt somit die unbekannten Funktionen A, B, C, D und dann aus (49) und (51) E, F, G, H. Dabei bleiben A vollkommen willkürlich, in B, C, D je eine Funktion von einer Veränderlichen, nämlich von y_2, in E eine willkürliche Konstante, F, G, H sind vollständig bestimmt.

Aus (48) folgen dann in bekannter Weise y_3, $\overline{y}_1$, $\overline{y}_2$, $\overline{y}_3$ als Funktionen von y_1 und y_2.

Ist $B_{y_2} = 0$ und $D_{y_2} \neq 0$, so erhält man statt (52) das System

$$A_{y_1} = \frac{2 A_{y_2} C_{y_2}}{D_{y_2}}, \quad B_{y_1} = A_{y_2}, \quad D_{y_1} = C_{y_2}.$$

Es ist aus demselben Grunde wie (52) verträglich. Jetzt bleibt C willkürlich, während A, B, D noch eine willkürliche Funktion von y_2 enthalten. Ist dagegen auch $D_{y_2} = 0$, so erhält man mit eventueller Vertauschung von y_3 und $\overline{y}_3$, was an der Allgemeinheit nichts ändert,

$$\left.\begin{aligned} A &= A(y_1) \\ B &= \text{konst.} \\ C &= C(y_1, y_2) \end{aligned}\right\} \text{willkürlich,}$$

D als Funktion von y_1 allein mit einer willkürlichen Konstanten.

[11]) E. Goursat, Leçons sur l'intégration des équations aux derivées partielles du premier ordre 1, S. 11, Paris 1891, A. Hermann.

Der Vollständigkeit halber sei noch erwähnt, daß die Integration der ifferentialgleichung (42) nach I, S. 282 durch Funktionen mit drei Para- etern auf lineare Räume führt, also die y als lineare Funktionen der y s allgemeines Integral liefert. Die Integration durch Funktionen von nem Parameter führt auf ametrische Kurven und gelingt in bekannter eise durch eine Quadratur. Integralmannigfaltigkeiten höherer Dimension s der dritten sind nach I, S. 281 ausgeschlossen.

Die Integration der entsprechenden Differentialgleichung in fünf Ver- nderlichen geschieht nach I, S. 281 entweder durch Kurven oder Flächen. iese sind nach S. 732 Torsen. Die Integration führt auf ein System von vei gewöhnlichen Differentialgleichungen in fünf unbekannten Funktionen [12]), bt also zu keinen Integrabilitätsbedingungen Anlaß. Bei vier Veränderlichen nd Integralkurven und -flächen möglich, diese sind aber nach I, S. 282 ır linear, also Ebenen. Hier treten sonach, ebenso wie bei drei und zwei eränderlichen, überhaupt keine Integrationsschwierigkeiten auf, es sind öchstens Quadraturen notwendig.

§ 6.

Eine spezielle ametrische Fläche des R_6 mit ∞^2 Tangentialebenen.

Es ist nun noch zu zeigen, daß es unter den im § 5 erhaltenen netrischen Flächen wirklich solche mit ∞^2 Tangentialebenen gibt. Wir eben eine solche Fläche an.

Setzt man $A=0$, $B=y_2$, $C=y_3$, $D=y$, $E=-\frac{y_2^2}{2}$, so erhält an ein partikuläres Integral des Systems (51), (52). Die entsprechende netrische Fläche hat die Gleichungen:

$$\begin{aligned} \bar{y}_1 &= -\frac{y_2^3}{6}, & \bar{y}_2 &= -\frac{y_1 y_2^2}{2}, \\ y_3 &= \frac{y_2^2}{2}, & \bar{y}_3 &= y_1 y_2 . \end{aligned} \tag{53}$$

e liegt in keinem linearen Raum niedrigerer Dimension, denn eine Beziehung on der Gestalt $c_1 y_1 + c_2 y_2 + c_3 y_3 + c_4 \bar{y}_1 + c_5 \bar{y}_2 + c_6 \bar{y}_3 + c_7 = 0$ mit onstanten c ist in y_1 und y_2 nur dann identisch erfüllt, wenn alle c ull sind. Die Fläche (53) hat wirklich ∞^2 Tangentialebenen. Denn die leichungen ihrer Tangentialebene im Punkte (y_1, y_2) lauten in den laufenden oordinaten $\eta_1, \eta_2, \eta_3, \bar{\eta}_1, \bar{\eta}_2, \bar{\eta}_3$:

$$\bar{\eta}_1 + \frac{y_2}{2}\eta_2 - \frac{y_2^3}{3} = 0, \qquad \bar{\eta}_2 + \frac{y_2^2}{2}\eta_1 + y_1 y_2 \eta_2 - y_1 y_2^2 = 0,$$

$$\eta_3 - y_2 \eta_2 + \frac{y_2^2}{2} = 0, \qquad \bar{\eta}_3 - y_2 \eta_1 - y_1 \eta_2 + y_1 y_2 = 0 .$$

[12]) Vgl. II, S. 31.

Diese Gleichungen hängen aber von den beiden Parametern y_1, y_2 wesentlich ab, d. h. nicht von einer funktionellen Kombination der beiden. Sie stellen daher ∞^2, nicht ∞^1 Tangentialebenen vor.

Um das einzusehen, gehen wir so vor: Wir betrachten die Tangentialebene im Punkte mit den bestimmten Parametern $y_1 = a$, $y_2 = b$, und in ihr den speziellen Punkt mit den Koordinaten $(\eta_1^0, \eta_2^0, \eta_3^0, \bar{\eta}_1^0, \bar{\eta}_2^0, \bar{\eta}_3^0)$. Wenn die Tangentialebene in einem anderen Flächenpunkt $(y_1 = a', y_2 = b')$ mit der im Punkte (a, b) zusammenfallen sollte, so müßten sich die zu den Werten a', b' gehörigen Koordinaten $\eta, \bar{\eta}$ so bestimmen lassen, daß sie mit den zu (a, b) gehörigen $\eta^0, \bar{\eta}^0$ übereinstimmen, und zwar für beliebige Werte von η_1^0 und η_2^0. Das liefert aber folgende Gleichungen:

$$\eta_1 = \eta_1^0, \qquad \eta_2 = \eta_2^0,$$

$$b'\eta_2 - \frac{b'^2}{2} = b\eta_2^0 - \frac{b^2}{2},$$

$$b'\eta_1 + a'\eta_2 - a'b' = b\eta_1^0 + a\eta_2^0 - ab,$$

daher, weil diese Gleichungen in η_1^0, η_2^0 identisch erfüllt sein sollen, $a = a'$, $b = b'$, d. h. wir erhalten zusammenfallende Tangentialebenen nur dann, wenn auch die betreffenden Berührungspunkte zusammenfallen.

(Eingegangen am 24. Dezember 1930.)

Über einen Satz von Vijayaraghavan.

Von

J. Karamata in Beograd.

Bei den Beweisen der Sätze von Tauberschem Charakter, wegen ihrer Vereinfachung, spielt eine wichtige Rolle der folgende Satz von H. T. Vijayaraghavan[1]):

Satz 1. *Sei* $f(r)=\sum_{\nu=0}^{\infty} u_\nu r^\nu$ *konvergent für* $r<1$ *und*

$$f(r)=O(1), \quad \text{bei} \quad r\to 1;$$

aus $\sum_{\nu=n}^{q} u_\nu \geqq -M$, *für alle* $n<q\leqq 3n$ *und* $n=1,2,3,\ldots$, *folgt*

$$\sum_{\nu=0}^{n} u_\nu = O(1), \quad \text{bei} \quad n\to\infty.$$

Da dieser Satz auch an und für sich ein gewisses Interesse darbietet, gebe ich hier einen einfachen Beweis desselben, indem ich ihn zugleich folgendermaßen verallgemeinere:

Satz 2. *Sei* $A(t)$ *eine Funktion von beschränkter Schwankung in jedem Intervall* $(0,T)$, *und existiere das Integral*

1) $$I(s)=\int_0^{\infty} e^{-st}\, d\{A(t)\} \quad \text{für jedes} \quad s>0;$$

aus

2) $$I(s)=O(1), \quad \text{bei} \quad s\to 0,$$

und

3) $$A(t)-A(x)\geqq -w(\lambda), \quad \text{für jedes} \quad x>0 \quad \text{und} \quad x<t\leqq x\lambda, \quad \lambda>1,$$

folgt

4) $$A(x)=O(1), \quad \text{bei} \quad x\to\infty.$$

[1]) T. Vijayaraghavan, A Tauberian theorem, Journal London Math. Soc. 1 (1926), 113–120.

Beweis. Ich betrachte die Funktion

$$\varphi(x) = (1+a)\,\bar{e}^{x} - a\,\bar{e}^{2x}, \quad \text{wo} \quad a > 1, \quad \text{und setze} \quad \lg\frac{2a}{1+a} = \alpha > 0,$$

dann ist $\varphi(0) = 1$, $\varphi(x)$ zunehmend für $0 < x < \alpha$ und abnehmend für $\alpha < x < \infty$.

Weiter ist, wenn

$$J(s) = (1+a)\,I(s) - a\,I(2s) \tag{5}$$

und $x = \frac{\alpha}{s}$ gesetzt wird,

$$J(s) - A(x) = \int_0^x \{\varphi(st) - 1\}\,d\{A(t)\} + \int_x^\infty \varphi(st)\,d\{A(t)\}$$
$$= \sum_{\nu=-\infty}^{0} \int_{x\lambda^{\nu-1}}^{x\lambda^{\nu}} \{\varphi(st) - 1\}\,d\{A(t)\} + \sum_{\nu=0}^{\infty} \int_{x\lambda^{\nu}}^{x\lambda^{\nu+1}} \varphi(st)\,d\{A(t)\},$$

und nach dem zweiten Mittelwertsatz

$$J(s) - A(x) = \sum_{\nu=-\infty}^{0} \{\varphi(\alpha\lambda^{\nu}) - 1\}\{A(x\lambda^{\nu}) - A(\xi_\nu')\}$$
$$+ \sum_{\nu=0}^{\infty} \varphi(\alpha\lambda^{\nu})\{A(\xi_\nu) - A(x\lambda^{\nu})\},$$

wo

$x\lambda^{\nu-1} \leqq \xi_\nu' \leqq x\lambda^{\nu}$, $\nu = 0, -1, -2, \ldots$ und $x\lambda^{\nu} \leqq \xi_\nu \leqq x\lambda^{\nu+1}$, $\nu = 0, 1, 2, \ldots$,

so daß nach (3)

$$J(s) - A(x) \geqq -w(\lambda)\Big[\sum_{\nu=-\infty}^{0} \{\varphi(\alpha\lambda^{\nu}) - 1\} + \sum_{\nu=0}^{\infty} \varphi(\alpha\lambda^{\nu})\Big].$$

Also wird

$$A(x) \leqq J\Big(\frac{\alpha}{x}\Big) + w(\lambda)\,M(\alpha, \lambda), \tag{6}$$

wo

$$M(\alpha, \lambda) = \sum_{\nu=-\infty}^{0} (1 - \bar{e}^{\alpha\lambda^{\nu}})(a\,\bar{e}^{\alpha\lambda^{\nu}} - 1) + \sum_{\nu=0}^{\infty} \bar{e}^{\alpha\lambda^{\nu}}(1 + a - a\,\bar{e}^{\alpha\lambda^{\nu}}),$$
$$\alpha = \lg\frac{2a}{1+a} \quad \text{und} \quad a > 1.$$

Da wegen (2) und (5)

$$J(s) = O(1), \quad \text{bei} \quad s \to 0,$$

ist, so folgt aus (6), daß

$$A(x) < O(1), \quad \text{bei} \quad x \to \infty. \tag{7}$$

Wird nun an Stelle von $\varphi(x)$ die folgende Funktion

$$\varphi(x) = \frac{a\,\bar{e}^{2x} - \bar{e}^{x}}{a-1}, \quad a > 1,$$

genommen, so liefern ähnliche Betrachtungen, daß

$$A(x) > O(1), \quad \text{bei} \quad x \to \infty. \tag{8}$$

Denn, wird $\lg 2a = \bar{\alpha}$ gesetzt, so ist $\bar{\varphi}(0) = 1$, $\bar{\varphi}(x)$ abnehmend wischen 0 und $\bar{\alpha}$, dann zunehmend und negativ.

Wird weiter $x = \frac{\bar{\alpha}}{s}$ und

9) $$\bar{J}(s) = \frac{a}{a-1} I(2s) - \frac{1}{a-1} I(s)$$

esetzt, so wird

$$\begin{aligned} I(x) - \bar{J}(s) &= \int_0^x \{1 - \bar{\varphi}(st)\}\, d\{A(t)\} + \int_x^\infty -\bar{\varphi}(st)\, d\{A(t)\} \\ &= \sum_{\nu=-\infty}^{0} \int_{x\lambda^{\nu-1}}^{x\lambda^{\nu}} \{1 - \bar{\varphi}(st)\}\, d\{A(t)\} + \sum_{\nu=0}^{\infty} \int_{x\lambda^{\nu}}^{x\lambda^{\nu+1}} -\bar{\varphi}(st)\, d\{A(t)\} \\ &= \sum_{\nu=-\infty}^{0} \{1 - \bar{\varphi}(\bar{\alpha}\lambda^{\nu})\}\{A(x\lambda^{\nu}) - A(\xi'_{\nu})\} \\ &\quad + \sum_{\nu=0}^{\infty} \{-\bar{\varphi}(\bar{\alpha}\lambda^{\nu})\}\{A(\xi_{\nu}) - A(x\lambda^{\nu})\} \\ &\geqq -w(\lambda)\Big[\sum_{\nu=-\infty}^{0} \{1 - \bar{\varphi}(\bar{\alpha}\lambda^{\nu})\} - \sum_{\nu=0}^{\infty} \bar{\varphi}(\bar{\alpha}\lambda^{\nu})\Big], \end{aligned}$$

o

$x\lambda^{\nu-1} \leqq \xi'_{\nu} \leqq x\lambda^{\nu}$, $\nu = 0, -1, -2, \ldots$ und $x\lambda^{\nu} \leqq \xi_{\nu} \leqq x\lambda^{\nu+1}$, $\nu = 0, 1, 2, \ldots$.

Also ist

10) $$A(x) > \bar{J}\left(\frac{\bar{\alpha}}{x}\right) - w(\lambda)\, \overline{M}(\bar{\alpha}, \lambda),$$

o

$$\overline{M}(\bar{\alpha}, \lambda) = \frac{1}{a-1}\Big[\sum_{\nu=-\infty}^{0} (1 - \bar{e}^{\bar{\alpha}\lambda^{\nu}})(a\,\bar{e}^{\bar{\alpha}\lambda^{\nu}} + a - 1) + \sum_{\nu=0}^{\infty} \bar{e}^{\bar{\alpha}\lambda^{\nu}}(1 - a\,\bar{e}^{\bar{\alpha}\lambda^{\nu}})\Big],$$

$$\bar{\alpha} = \lg 2a \quad \text{und} \quad a > 1.$$

Wegen (2) und (9) folgt nun (8) aus (10), und aus (8) und (7) die ehauptung (4).

Zum Schluß sei noch folgendes bemerkt:

Erstens, daß der vorstehende Beweis fast ungeändert verwendet werden ann, um den folgenden allgemeineren Satz zu beweisen:

Satz 3. *Sei $A(t)$ von beschränkter Schwankung in jedem Intervall $(0, T)$, und existiere das Integral*

$$I(s) = \int_0^\infty \bar{e}^{st}\, d\{A(t)\} \quad \textit{für jedes} \quad s > 0,$$

ei ferner $g(x)$ stetig und positiv in $(0, \infty)$ und

$$g(2x) = O\{g(x)\}, \quad \textit{bei} \quad x \to \infty;$$

us

1) $$I(s) = O\left\{g\left(\frac{1}{s}\right)\right\}, \quad \textit{bei} \quad s \to 0,$$

und

(12) $A(t)-A(x)\geqq -w(\lambda)g(x)$, *für jedes* $x>0$ *und* $x<t\leqq x\lambda$, $\lambda>1$,

folgt dann

(13) $$A(x)=O\{g(x)\}, \quad \textit{bei} \quad x\to\infty.$$ [2])

Zweitens daß dasselbe Prinzip, indem man die Funktionen $\varphi(x)$ und $\overline{\varphi}(x)$ durch passend gewählte Polynome von $\overline{e}^{x}$ ersetzt, auch dazu verwendet werden kann, um die Tauberschen Sätze (besonders Satz A, S. 33, meiner unter [2]) zitierten Arbeit) direkt zu beweisen, wie ich es unlängst in einer ganz allgemeinen Form gezeigt habe[3]).

Beograd, den 25. Februar 1931.

---

[2]) Würde in diesem Satze statt (13) nur verlangt, daß $A(x)>O\{g(x)\}$ wird, so verliert er für zu schnell wachsende Funktionen $g(x)$ {z. B. wenn $g(x)\sim x^{\sigma}$, $\sigma>0$} seinen Sinn, da in diesem Falle $A(x)>O\{g(x)\}$ schon aus (12) folgt, wie ich es in meiner Arbeit: J. Karamata, Neuer Beweis und Verallgemeinerung der Tauberschen Sätze, welche die Laplaceschen Stieltjesschen Transformationen betreffen, Journal für die reine u. angewandte Math. **164** (1931), S. 36 gezeigt habe.

Zugleich sei bemerkt, daß der Satz 3 den auf S. 36 meiner eben zitierten Arbeit angeführten Hilfssatz enthält, wodurch die Beweise dieser Arbeit etwas verkürzt werden, und in welchem Falle die Beweisanordnung diejenige des Herrn Vijayaraghavan wäre, welche er (siehe loc. cit. [1])) für Potenzreihen verwendet hat.

[3]) J. Karamata, Sur le rapport entre les convergences d'une suite de fonctions et de leurs moments, et son application à l'inversion des procédés de sommabilité, Studia Mathematica **3** (1931), S. 68–76, Lwów.

(Eingegangen am 28. Februar 1931.)

Verschärfung eines Harnackschen Satzes und anderer Abschätzungen für nichtnegative harmonische Polynome.

Von

E. Egerváry in Budapest.

§ 1.

Es sei $f(r,t)$ eine im Einheitskreise $r<1$ reguläre, nichtnegative, harmonische Funktion, welche für $r=0$ den Wert 1 annimmt, mit der für $r<1$ konvergenten Entwicklung

$$f(r,t)=1+\sum_{\nu=1}^{\infty} r^{\nu}(a_{\nu}\cos\nu t+b_{\nu}\sin\nu t)>0 \quad \text{für} \quad r<1,$$

dann ist nach einem Satze von Harnack

$$\frac{1-r}{1+r}\leqq f(r,t)\leqq\frac{1+r}{1-r};$$

und Gleichheit gilt in (2) für die Funktion

$$f^{*}(r,t)=\frac{1-r^{2}}{1-2r\cos t+r^{2}}$$

(links für $t=\pi$, rechts für $t=0$).

Für die Gesamtheit der Funktionen (1) gelten noch die Abschätzungen

$$\left(\frac{1-r}{1+r}\right)^{2}\leqq\frac{f(r_1,t_1)}{f(r_2,t_2)}\leqq\left(\frac{1+r}{1-r}\right)^{2} \qquad (r_1,r_2\leqq r<1),$$

$$|\operatorname{grad} f(r,t)|=\sqrt{\left(\frac{\partial f}{\partial r}\right)^{2}+\left(\frac{1}{r}\frac{\partial f}{\partial t}\right)^{2}}\leqq\frac{2}{(1-r)^{2}} \qquad (r<1),$$

$$\sqrt{a_{\nu}^{2}+b_{\nu}^{2}}\leqq 2 \qquad (\nu=1,2,\ldots),$$

und das Gleichheitszeichen in (4) und (5) tritt bei der Funktion (3) auf, während die Entwicklungskoeffizienten von (3) für $r\to 1$ der oberen Grenze 2 (6) beliebig nahekommen.

Betrachtet man statt der Gesamtheit (1) die Gesamtheit der im Einheitskreise $r<1$ nichtnegativen, harmonischen Polynome höchstens n-ter

Ordnung, welche für $r=0$ den Wert 1 annehmen, also die Funktionen

$$(7) \qquad P_n(r,t)=1+\sum_{\nu=1}^{n} r^\nu(a_\nu\cos\nu t+b_\nu\sin\nu t)>0 \quad \text{für} \quad r<1,$$

so lassen sich die obigen Ungleichungen verschärfen. Für solche Polynome ist die Ungleichung (2) für $r=1$ sowie die Ungleichung (6) von Herrn L. Fejér[1]), die aus (5) folgende Ungleichung $\left|\frac{\partial f}{\partial t}\right| \leqq \frac{2r}{(1-r)^2}$ für $r=1$ von Herrn O. Szász und vom Verfasser[2]) verschärft worden.

In vorliegender Arbeit gebe ich zunächst die Verschärfung der Harnackschen Ungleichung (2) für beliebiges r $(0 \leqq r \leqq 1)$, und zwar in folgender Form:

Ist

$$(7) \qquad P_n(r,t)=1+\sum_{\nu=1}^{n} r^\nu(a_\nu\cos\nu t+b_\nu\sin\nu t)>0 \quad \text{für} \quad r<1,$$

so ist einerseits

$$(8) \qquad P_n(r,t) \geqq \frac{\sin(n+1)\theta_M}{\sin\theta_M} \quad \textit{für} \quad r \leqq \frac{\cos(n+2)\frac{\theta_M}{2}}{\cos n\frac{\theta_M}{2}} \qquad \left(0 \leqq \theta_M \leqq \frac{\pi}{n+2}\right)$$

andererseits

$$(9) \qquad P_n(r,t) \geqq \frac{\sin(n+1)\vartheta_m}{\sin\vartheta_m} \quad \textit{für} \quad r \leqq \left|\frac{\cos(n+2)\frac{\vartheta_m}{2}}{\cos n\frac{\vartheta_m}{2}}\right| \qquad \left(\frac{\pi}{n+2} \leqq \vartheta_m \leqq \frac{\pi}{n+1}\right)$$

und Gleichheit gilt nur für das harmonische Polynom[3])

$$(10) \qquad P_n^*(r,t)=1+\frac{2}{\frac{\sin(n+1)\theta}{\sin\theta}+n+1}\sum_{\nu=1}^{n}\left[\frac{\sin(n-\nu+1)\theta}{\sin\theta}+(n-\nu+1)\cos\nu\theta\right]r$$

(in (8) für $\theta=\theta_M$, in (9) für $\theta=\theta_m=\pi-\vartheta_m$), *welches am Kreisrand* $r=1$ *in das Kosinuspolynom n-ter Ordnung*

$$(11) \qquad P_n^*(1,t)=\frac{1}{2\left[\frac{\sin(n+1)\theta}{\sin\theta}+n+1\right]}\left[\frac{\sin(n+1)\frac{t+\theta}{2}}{\sin\frac{t+\theta}{2}}+\frac{\sin(n+1)\frac{t-\theta}{2}}{\sin\frac{t-\theta}{2}}\right]^2$$

übergeht.

---

[1]) L. Fejér, a) Über trigonometrische Polynome, Journal für die reine und angewandte Mathematik 146 (1915), S. 53–82; b) Über eine Aufgabe der Harnackschen Potentialtheorie, Gött. Nachr. 1928, S. 109–117.

[2]) E. Egerváry und O. Szász, Einige Extremalprobleme im Bereiche der nichtnegativen trigonometrischen Polynome, Math. Zeitschr. 27 (1928), S. 641–652.

[3]) Oder für $P_n^*(r, t-t_0)$.

Für $\theta_M = 0$ geht die Ungleichung (8) in die Fejérsche Ungleichung $P_n(r,t) \leqq n+1$ für $r \leqq 1$ über, während (10) und (11) in die Polynome

$$(r,t) = 1 + \frac{2}{n+1}\sum_{\nu=1}^{n}(n-\nu+1)r^\nu \cos \nu t \quad \text{bzw.} \quad P_n^*(1,t) = \frac{1}{n+1}\left(\frac{\sin(n+1)\frac{t}{2}}{\sin\frac{t}{2}}\right)^2$$

übergehen. Für $\theta_M = \frac{\pi}{n+2}$ geht $P_n^*(r,t)$ in das Polynom

$$P_n^*(r,t) = 1 + 2\cos\frac{\pi}{n+2} r\cos t + \ldots$$

über, bei welchem für die erste Amplitude $\sqrt{a_1^2+b_1^2}$ der bekannte Fejérsche Höchstwert $2\cos\frac{\pi}{n+2}$ auftritt, d. h. es ist

$$\operatorname{Max}\left(\frac{\partial P_n^*(r,t)}{\partial r}\right)_{r=0} = \sqrt{a_1^2+b_1^2} = 2\cos\frac{\pi}{n+2}.$$

Die beiden obenerwähnten Sätze des Extremums von L. Fejér sind hiernach in meinem allgemeinen Resultat ((8)—(11)) enthalten.

Nachher befasse ich mich mit dem noch etwas allgemeineren Extremumproblem, bei welchem der Wert des nichtnegativen harmonischen Polynoms nicht, wie bis jetzt, an der Stelle $r = 0$, sondern an einer beliebigen Stelle der Einheitskreisfläche vorgeschrieben ist. Von den gewonnenen Ergebnissen sei hier folgende Verschärfung der Ungleichungen (4) hervorgehoben:

Ist

$$P_n(r,t) = a_0 + \sum_{\nu=1}^{n} r^\nu (a_\nu \cos \nu t + b_\nu \sin \nu t) > 0 \quad \text{für} \quad r < 1,$$

so ist für

$$(12) \qquad 0 \leqq t_1, t_2 \leqq 2\pi, \qquad 0 \leqq r_1, r_2 \leqq \sqrt{\frac{\cos(n+2)\frac{\theta}{2}}{\cos(n-2)\frac{\theta}{2}}}, \qquad 0 \leqq \theta \leqq \frac{\pi}{n+2},$$

$$(13) \qquad \frac{1}{\lambda_\theta} \leqq \frac{P_n(r_1,t_1)}{P_n(r_2,t_2)} \leqq \lambda_\theta, \qquad \lambda_\theta = \left[\frac{\cos n\frac{\theta}{2} + \sqrt{\cos(n+2)\frac{\theta}{2}\cos(n-2)\frac{\theta}{2}}}{\sin\theta}\right]^2,$$

und Gleichheit gilt nur für das harmonische Polynom $P_n^*(r,t)$,[4]) welches am Kreisrande $r = 1$ in das Kosinuspolynom n-ter Ordnung

$$(14) \qquad P_n^*(1,t) = \left[\frac{\sin n\frac{t+\theta}{2}}{\sin\frac{t+\theta}{2}}\cos\frac{t-\theta}{2} + \frac{\sin n\frac{t-\theta}{2}}{\sin\frac{t-\theta}{2}}\cos\frac{t+\theta}{2}\right]^2$$

übergeht.

[4]) Oder für konst. $P_n^*(r, t-t_0)$.

Die vorher erwähnte Fejérsche Frage läßt sich auch so fassen: Fü ein im Einheitskreise $r \leqq 1$ nichtnegatives, harmonisches Polynom n-ter Ord nung ist der Wert, den es an einer gegebenen Stelle des Einheitskreis randes $r = 1$ annimmt, gegeben. Gesucht wird das Minimum des Polynom im Mittelpunkte $r = 0$. In einer unlängst erschienenen Arbeit[5]) löst Her Serge Bernstein mit einer bemerkenswerten neuen Methode — unter anderen — das Problem, welches entsteht, wenn man in zwei gegebenen Punkten de Kreisrandes $r = 1$ den Wert des Polynoms vorschreibt und wieder da Minimum im Mittelpunkte $r = 0$ sucht. Im folgenden bestimme ich nur das Minimum in einem beliebigen vorgeschriebenen Punkte des Kreisinnern und zwar im wesentlichen auf Grund der älteren Fejérschen Methode. Da verallgemeinerte S. Bernsteinsche Resultat lautet:

Ist der Wert eines nichtnegativen harmonischen Polynoms in zwe Randpunkten des Einheitskreises $r = 1$ vorgeschrieben:

$$P_n(1, \varphi) = P_\varphi \geqq 0, \qquad P_n(1, \psi) = P_\psi \geqq 0,$$

so ist (für $\varrho < 1$)

$$(15) \qquad P_n(\varrho, 0) \geqq \frac{P_\varphi S_n(0, \psi) + P_\psi S_n(0, \varphi) - 2\left|\sqrt{P_\varphi P_\psi}\right| S_n\left(\frac{\varphi - \psi}{2}, \varphi\right)}{S_n(0, \varphi) S_n(0, \psi) - S_n\left(\frac{\varphi - \psi}{2}, \varphi\right)^2},$$

wobei

$$S_n(\alpha, \beta) = \frac{\sin(n+1)\alpha - 2\varrho \sin n\alpha \cos(\alpha - \beta) + \varrho^2 \sin(n-1)\alpha}{(1 - \varrho^2) \sin \alpha}$$

bezeichnet, und Gleichheit gilt in (15) nur für das harmonische Polynom $P_n^*(r, t)$, dessen Randpolynom $P_n^*(1, t)$ sich aus der Gleichung

$$(16) \qquad \begin{vmatrix} \sqrt{P_n^*(1,t)} & \sqrt{P_\varphi} & \sqrt{P_\psi} \\ S_n\left(\frac{\varphi - t}{2}, \varphi\right) & S_n(0, \varphi) & S_n\left(\frac{\varphi - \psi}{2}, \varphi\right) \\ S_n\left(\frac{\psi - t}{2}, \psi\right) & S_n\left(\frac{\psi - \varphi}{2}, \psi\right) & S_n(0, \psi) \end{vmatrix} = 0$$

unmittelbar ergibt.

Für $\varrho = 0$ geht nun tatsächlich die rechte Seite der Ungleichung (15) in den Bernsteinschen Wert des Minimums über.

Durch den Grenzübergang $\varphi \to \psi$ gelange ich hieraus zu einer Abschätzung der Derivierten $\frac{\partial P_n(1, t)}{\partial t}$, welche eine Verallgemeinerung eines Resultates von Herrn O. Szász und von mir[6]) darstellt.

---

[5]) S. Bernstein, Sur l'application de la méthode de Tchebycheff à une classe de problèmes de M. Fejér, Bulletin de l'Académie de l'Urss 1930, S. 381–398.

[6]) loc. cit. [2]), S. 650–652.

§ 2.

Bei der Lösung aller vorher genannten Fragen dient als Grundlage die Fejérsche Parameterdarstellung der nichtnegativen, trigonometrischen Polynome[7]), welche (in einer etwas verallgemeinerten Fassung) folgendermaßen lautet:

Gegeben seien $2n+1$ reelle Zahlen $\alpha_0, \alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n$; die Zahlenmenge

(*) $$L_P = \alpha_0 a_0 + \alpha_1 a_1 + \beta_1 b_1 + \ldots + \alpha_n a_n + \beta_n b_n$$

— wobei $P_n(r,t) = a_0 + \sum_{\nu=1}^{n} r^\nu (a_\nu \cos \nu t + b_\nu \sin \nu t)$ die Gesamtheit derjenigen nichtnegativen, harmonischen Polynome durchläuft, welche für $r = \mathsf{P} < 1$, $t = 0$ den Wert 1 annehmen — ist identisch mit dem Wertevorrat der Hermiteschen Form

$$H_n(x) = \sum_{\mu=0}^{n} \sum_{\nu=0}^{n} \gamma_{\mu-\nu} x_\nu \bar{x}_\mu \qquad (\gamma_\nu = \alpha_\nu + i\beta_\nu;\ \gamma_{-\nu} = \alpha_\nu - i\beta_\nu)$$

für

$$\sum_{\mu=0}^{n} \sum_{\nu=0}^{n} \mathsf{P}^{|\mu-\nu|} x_\nu \bar{x}_\mu = 1.$$

Bezeichnen daher λ_m bzw. λ_M die kleinste bzw. größte Wurzel der charakteristischen Gleichung

$$\begin{vmatrix} \alpha_0 - \lambda & \alpha_1 + i\beta_1 - \mathsf{P}\lambda & \ldots & \alpha_n + i\beta_n - \mathsf{P}^n\lambda \\ \alpha_1 - i\beta_1 - \mathsf{P}\lambda & \alpha_0 - \lambda & \ldots & \alpha_{n-1} + i\beta_{n-1} - \mathsf{P}^{n-1}\lambda \\ \ldots & \ldots & \ldots & \ldots \\ \alpha_n - i\beta_n - \mathsf{P}^n\lambda & \alpha_{n-1} - i\beta_{n-1} - \mathsf{P}^{n-1}\lambda & \ldots & \alpha_0 - \lambda \end{vmatrix} = 0,$$

so ist offenbar

$$\lambda_m \leqq L_P \leqq \lambda_M.$$

Die harmonischen Polynome $P_n^*(r,t)$, für die hier rechts (bzw. links) Gleichheit gilt, werden — nach Auflösung des homogenen, linearen Gleichungssystems

$$\begin{aligned} &(\alpha_0 - \lambda)x_0 + (\alpha_1 + i\beta_1 - \mathsf{P}\lambda)x_1 + \ldots + (\alpha_n + i\beta_n - \mathsf{P}^n\lambda)x_n = 0, \\ &(\alpha_1 - i\beta_1 - \mathsf{P}\lambda)x_0 + (\alpha_0 - \lambda)x_1 + \ldots + (\alpha_{n-1} + i\beta_{n-1} - \mathsf{P}^{n-1}\lambda)x_n = 0, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ &(\alpha_n - i\beta_n - \mathsf{P}^n\lambda)x_0 + (\alpha_{n-1} - i\beta_{n-1} - \mathsf{P}^{n-1}\lambda)x_1 + \ldots + (\alpha_0 - \lambda)x_n = 0 \end{aligned}$$

$(\lambda = \lambda_M$ (bzw. $\lambda_m))$

---

[7]) L. Fejér, loc. cit. [1]) a); vgl. O. Szász, Über harmonische Funktionen und L-Formen, Math. Zeitschr. 1 (1918), S. 149–162.

mit der Nebenbedingung

$$\sum_{\mu=0}^{n}\sum_{\nu=0}^{n} \mathsf{P}^{|\mu-\nu|}\, x_\nu \bar{x}_\mu = 1$$

— bestimmt durch

$$P_n^*(1,t) = |\, x_0 + x_1 z + \ldots + x_n z^n \,|^2_{z=e^{it}}.$$

§ 3.

Wir betrachten zunächst die Gesamtheit derjenigen nichtnegativen harmonischen Polynome n-ter Ordnung, welche für $r=0$ den Wert 1 annehmen:

$$P_n(r,t) = 1 + \sum_{\nu=1}^{n} r^\nu (a_\nu \cos \nu t + b_\nu \sin \nu t) > 0 \quad \text{für} \quad r<1,$$

und fragen nach den extremen Werten von $P_n(\varrho, 0)$ für eine beliebige[8]), aber festgewählte innere Stelle $(\varrho, 0)$ des Einheitskreises.

Zu diesem Zwecke haben wir nach § 2 zu setzen:

$$\mathsf{P}=0, \quad \alpha_\nu = \varrho^\nu; \quad \beta_\nu = 0 \qquad (\nu = 0, 1, 2, \ldots, n, \ 0<\varrho<1);$$

zu bestimmen sind also die extremen Werte der quadratischen Form

$$P_n(\varrho, 0) = \sum_{\mu=0}^{n}\sum_{\nu=0}^{n} \varrho^{|\mu-\nu|}\, x_\mu x_\nu \tag{17}$$

mit der Nebenbedingung

$$P_n(0,0) = \sum_{\nu=0}^{n} x_\nu^2 = 1. \tag{17'}$$

Durch Nullsetzen der partiellen Derivierten $L_k = \frac{\partial}{\partial x_k}[P_n(\varrho,0) - \lambda P_n(0,0)]$ erhält man das Gleichungssystem

$$\begin{aligned}
&L_0 \equiv (1-\lambda)x_0 + \varrho x_1 + \varrho^2 x_2 + \ldots + \varrho^n x_n = 0,\\
&L_1 \equiv \varrho x_0 + (1-\lambda)x_1 + \varrho x_2 + \ldots + \varrho^{n-1} x_n = 0,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&L_n \equiv \varrho^n x_0 + \varrho^{n-1} x_1 + \varrho^{n-2} x_2 + \ldots + (1-\lambda)x_n = 0
\end{aligned} \tag{18}$$

mit der charakteristischen Gleichung

$$D_{n+1}(\lambda;\varrho) = \begin{vmatrix}
1-\lambda & \varrho & \varrho^2 & \ldots & \varrho^n\\
\varrho & 1-\lambda & \varrho & \ldots & \varrho^{n-1}\\
\varrho^2 & \varrho & 1-\lambda & \ldots & \varrho^{n-2}\\
\ldots & \ldots & \ldots & \ldots & \ldots\\
\varrho^n & \varrho^{n-1} & \varrho^{n-2} & \ldots & 1-\lambda
\end{vmatrix} = 0. \tag{19}$$

[8]) Man kann offenbar ohne Einschränkung der Allgemeinheit $t=0$ annehmen.

Wir ersetzen die Gleichungen durch folgendes äquivalentes System[9]):

$$-\frac{1}{\varrho}L_0+L_1\equiv\left[\left(\varrho-\frac{1}{\varrho}\right)+\frac{\lambda}{\varrho}\right]x_0-\lambda x_1=0,$$

$$x_{k-1}-\left(\varrho+\frac{1}{\varrho}\right)L_k+L_{k+1}\equiv-\lambda x_{k-1}+\left[\left(\varrho-\frac{1}{\varrho}\right)+\left(\varrho+\frac{1}{\varrho}\right)\lambda\right]x_k-\lambda x_{k+1}=0,$$

$$L_{n-1}-\frac{1}{\varrho}L_n\equiv-\lambda x_{n-1}+\left[\left(\varrho-\frac{1}{\varrho}\right)+\frac{\lambda}{\varrho}\right]x_n=0$$

$$(k=1,2,\ldots,n-1).$$

Die entsprechende charakteristische Gleichung ist

$$\frac{1-\varrho^2}{\varrho^{n+1}}D_{n+1}(\lambda;\varrho)$$

$$\begin{vmatrix}
\left(\varrho-\frac{1}{\varrho}\right)\frac{1}{\lambda}+\frac{1}{\varrho} & -1 & 0 & \ldots & 0\\
-1 & \left(\varrho-\frac{1}{\varrho}\right)\frac{1}{\lambda}+\varrho+\frac{1}{\varrho} & -1 & \ldots & 0\\
0 & -1 & \left(\varrho-\frac{1}{\varrho}\right)\frac{1}{\lambda}+\varrho+\frac{1}{\varrho} & \ldots & 0\\
\ldots & \ldots & \ldots & \ldots & \ldots\\
0 & 0 & 0 & \ldots & \left(\varrho-\frac{1}{\varrho}\right)\frac{1}{\lambda}+\frac{1}{\varrho}
\end{vmatrix}$$

$$=\lambda^{n+1}\frac{\sin(n+2)\theta-2\varrho\sin(n+1)\theta+\varrho^2\sin n\theta}{\sin\theta}=0,\ ^{10})$$

θ durch die Gleichung

$$2\cos\theta=\left(\varrho-\frac{1}{\varrho}\right)\frac{1}{\lambda}+\varrho+\frac{1}{\varrho}$$

bestimmt ist. Es ist also

$$\lambda=\frac{1-\varrho^2}{1-2\varrho\cos\theta+\varrho^2}.$$

Die Gleichung (21) in θ besitzt $2n+2$ reelle, einfache, von 0 und π verschiedene Wurzeln[11]), und die größte Wurzel λ_M von (19) entspricht

---

[9]) Die hiermit eingeführten Linearformen sind unabhängig, weil die Substitutionsdeterminante $\frac{1-\varrho^2}{\varrho^{n+1}}$ für $0<\varrho<1$ nicht verschwindet.

[10]) Ergibt sich unmittelbar aus der bekannten Formel

$$\left\{\begin{matrix}n\text{ Zeilen}\end{matrix}\right\}\begin{vmatrix}
2\cos\theta & -1 & 0 & \ldots & 0\\
-1 & 2\cos\theta & -1 & \ldots & 0\\
0 & -1 & 2\cos\theta & \ldots & 0\\
\ldots & \ldots & \ldots & \ldots & \ldots\\
0 & 0 & 0 & \ldots & 2\cos\theta
\end{vmatrix}=\frac{\sin(n+1)\theta}{\sin\theta}$$

unter Beachtung der Entwicklungsregel für geränderte Determinanten.

[11]) Unmittelbar ersichtlich aus der Form (24) von $D_{n+1}(\lambda;\varrho)$.

nach (23) offenbar der zu 0 am nächsten liegenden Wurzel θ_M von (21
Die Gleichung (21) läßt sich aber folgendermaßen schreiben:

$$(24)\quad D_{n+1}(\lambda;\varrho)\doteq\frac{\left[\cos(n+2)\frac{\theta}{2}-\varrho\cos n\frac{\theta}{2}\right]\left[\sin(n+2)\frac{\theta}{2}-\varrho\sin n\frac{\theta}{2}\right]}{\cos\frac{\theta}{2}\cdot\sin\frac{\theta}{2}}=0$$

Der erste Faktor rechts verschwindet (bei der Annahme $0<\varrho<1$) genau einmal, der zweite überhaupt nicht im Intervalle $\left(0,\frac{\pi}{n+2}\right)$. Folglich ist $0<\theta_M<\frac{\pi}{n+2}$ und

$$(25)\qquad \cos(n+2)\frac{\theta_M}{2}-\varrho\cos n\frac{\theta_M}{2}=0\,.$$

Man kann also sowohl die Größe ϱ als auch die zugehörige größte Wurzel λ_M als Funktionen des Parameters θ_M ausdrücken. Nach (25) ist nämlich

$$(26)\qquad \varrho=\frac{\cos(n+2)\frac{\theta_M}{2}}{\cos n\frac{\theta_M}{2}}\qquad\left(0<\theta_M<\frac{\pi}{n+2}\right)$$

also nach (23)

$$(27)\qquad \lambda=\frac{1-\left[\frac{\cos(n+2)\frac{\theta_M}{2}}{\cos n\frac{\theta_M}{2}}\right]^2}{1-2\cos\theta_M\frac{\cos(n+2)\frac{\theta_M}{2}}{\cos n\frac{\theta_M}{2}}+\left[\frac{\cos(n+2)\frac{\theta_M}{2}}{\cos n\frac{\theta_M}{2}}\right]^2}=\frac{\sin(n+1)\theta_M}{\sin\theta_M}$$

$$\left(0<\theta_M<\frac{\pi}{n+2}\right).$$

Durchläuft der Parameter das Intervall $\left(0,\frac{\pi}{n+2}\right)$, so durchläuft ϱ das Intervall $(0,1)$, und die entsprechenden Werte von λ_M sind durch den Ausdruck (27) gegeben.

Wir haben also bewiesen:

Satz I. *Ist das harmonische Polynom*

$$P_n(r,t)=1+\sum_{\nu=1}^{n}r^\nu(a_\nu\cos\nu t+b_\nu\sin\nu t)>0\quad \textit{für}\quad r<1,$$

so ist für

$$(28)\qquad 0\leqq r\leqq\frac{\cos(n+2)\frac{\theta_M}{2}}{\cos n\frac{\theta_M}{2}}\qquad\left(0\leqq\theta_M\leqq\frac{\pi}{n+2}\right)$$

$$(29')\qquad P_n(r,0)=1+\sum_{\nu=1}^{n}a_\nu r^\nu\leqq\frac{\sin(n+1)\theta_M}{\sin\theta_M}\,;$$

so auch allgemein

9) $$P_n(r,t) \leqq \frac{\sin(n+1)\theta_M}{\sin\theta_M}.$$

Die kleinste Wurzel λ_m entspricht nach (23) der zu π am nächsten genden Wurzel θ_m von (21). Setzt man dort $\theta = \pi - \vartheta$, so folgt

0) $$D_{n+1}(\lambda;\varrho) \doteq \frac{\left[\cos(n+2)\frac{\vartheta}{2} + \varrho\cos n\frac{\vartheta}{2}\right]\left[\sin(n+2)\frac{\vartheta}{2} + \varrho\sin n\frac{\vartheta}{2}\right]}{\cos\frac{\vartheta}{2}\cdot\sin\frac{\vartheta}{2}}.$$

er erste Faktor rechts ist (bei der Annahme $0<\varrho<1$) im Inter- ll $\left(0, \frac{\pi}{n+2}\right)$ von 0 verschieden, verschwindet im Intervall $\left(\frac{\pi}{n+2}, \frac{\pi}{n+1}\right)$ nau einmal, während der zweite Faktor im ganzen Intervall $\left(0, \frac{\pi}{n+1}\right)$ nicht rschwindet. Demzufolge ist $\frac{\pi}{n+2} < \vartheta_m < \frac{\pi}{n+1}$, d. h. $\frac{n\pi}{n+1} < \theta_m < \frac{(n+1)\pi}{n+2}$ d

1) $$\cos(n+2)\frac{\pi-\theta_m}{2} + \varrho\cos n\frac{\pi-\theta_m}{2} = 0.$$

an kann also die Größe ϱ und die zugehörige kleinste Wurzel λ_m als nktionen des Parameters θ_m folgendermaßen ausdrücken:

2) $$\varrho = \left|\frac{\cos(n+2)\frac{\pi-\theta_m}{2}}{\cos n\frac{\pi-\theta_m}{2}}\right|, \qquad \lambda_m = \left|\frac{\sin(n+1)(\pi-\theta_m)}{\sin(\pi-\theta_m)}\right|$$

$$\left(\frac{n\pi}{n+1} < \theta_m < \frac{(n+1)\pi}{n+2}\right).$$

gilt also der

Satz I′. *Ist das harmonische Polynom*

$$P_n(r,t) = 1 + \sum_{\nu=1}^{n} r^\nu(a_\nu\cos\nu t + b_\nu\sin\nu t) > 0 \quad \textit{für} \quad r<1,$$

ist für

3) $$0 \leqq r \leqq \left|\frac{\cos(n+2)\frac{\pi-\theta_m}{2}}{\cos n\frac{\pi-\theta_m}{2}}\right| \qquad \left(\frac{n\pi}{n+1} < \theta_m < \frac{(n+1)\pi}{n+2}\right)$$

4) $$\left|\frac{\sin(n+1)\theta_m}{\sin\theta_m}\right| \leqq P_n(r,t).$$

Um dasjenige Polynom, für welches in (29) bzw. in (34) Gleichheit t, zu bestimmen, hat man noch die Unbekannten $x_0, x_1, \ldots, x_n$ aus dem eichungssystem (18) oder (20) zu berechnen. Die zur ersten Zeile von $_{n+1}$ gehörigen Unterdeterminanten sind

$$D_{n+1}^{(0,\nu)} \doteq \frac{\sin(\nu+1)\theta - \varrho\sin\nu\theta}{\sin\theta} \qquad (\nu = 0, 1, 2, \ldots, n);$$

also für $\varrho = \dfrac{\cos(n+2)\dfrac{\theta}{2}}{\cos n \dfrac{\theta}{2}}$

$$x_\nu \doteq D_{n+1}^{(0,\nu)} \doteq \cos\left(\frac{n}{2} - \nu\right)\theta. \tag{35}$$

Das (im wesentlichen) einzige harmonische Polynom $P_n^*(r,t)$, für welches in (29) bzw. in (34) Gleichheit gilt, ist also für $r=1$

$$P_n^*(1,t) = \frac{\left|\cos\frac{n}{2}\theta + \cos\left(\frac{n}{2}-1\right)\theta\cdot z + \ldots + \cos\left(\frac{n}{2}-n\right)\theta\cdot z^n\right|_{z=e^{it}}^2}{\cos^2\frac{n}{2}\theta + \cos^2\left(\frac{n}{2}-1\right)\theta + \ldots + \cos^2\left(\frac{n}{2}-n\right)\theta} \tag{36}$$

$$= \frac{1}{2\left[\dfrac{\sin(n+1)\theta}{\sin\theta} + n + 1\right]}\left[\frac{\sin(n+1)\dfrac{t+\theta}{2}}{\sin\dfrac{t+\theta}{2}} + \frac{\sin(n+1)\dfrac{t-\theta}{2}}{\sin\dfrac{t-\theta}{2}}\right]^2$$

$$(\theta = \theta_M \text{ bzw. } \theta = \theta_m),$$

für $r \leqq 1$

$$P_n^*(r,t) = 1 + \frac{2}{\dfrac{\sin(n+1)\theta}{\sin\theta} + n + 1} - \sum_{\nu=1}^{n} r^\nu\left[\frac{\sin(n-\nu+1)\theta}{\sin\theta} + (n-\nu+1)\cos\nu\theta\right] \tag{37}$$

$$(\theta = \theta_M \text{ bzw. } \theta = \theta_m).$$

Die Ungleichung (29) geht für $\theta_M \to 0$, d. h. für $\varrho \to 1$ in die bekannte Fejérsche Ungleichung

$$P_n(r,t) \leqq n+1 \quad \text{für} \quad 0 \leqq r \leqq 1,$$

der Ausdruck (36) in das entsprechende Extremalpolynom

$$P_n^*(1,t) = \frac{1}{n+1}\left(\frac{\sin(n+1)\dfrac{t}{2}}{\sin\dfrac{t}{2}}\right)^2$$

über.

Aus (26) und (29) erhält man

$$\frac{P_n(\varrho,t) - P_n(0,t)}{\varrho} \leqq \left[\frac{\sin(n+1)\theta_M}{\sin\theta_M} - 1\right]\frac{\cos n\dfrac{\theta_M}{2}}{\cos(n+2)\dfrac{\theta_M}{2}} = \frac{\sin n\theta_M}{\sin\theta_M},$$

also für $\varrho \to 0$, d. h. für $\theta_M \to \dfrac{\pi}{n+2}$

$$\left[\frac{\partial P_n(r,t)}{\partial r}\right]_{r=0} = a_1\cos t + b_1\sin t \leqq \frac{\sin n\dfrac{\pi}{n+2}}{\sin\dfrac{\pi}{n+2}} = 2\cos\frac{\pi}{n+2}$$

nd, da diese Ungleichung für jedes t gilt,

$$\sqrt{a_1^2 + b_1^2} \leqq 2\cos\frac{\pi}{n+2}.$$

as ist die verschärfte Ungleichung (6) für die erste Amplitude[12]) eines nicht-egativen harmonischen Polynoms, welche ebenfalls zuerst von Herrn Fejér ngegeben worden ist. Das entsprechende Extremalpolynom ist nach (37)

$$P_n^*(r,t) = 1 + 2\cos\frac{\pi}{n+2}\cdot r\cdot\cos t + \dots .$$

§ 4.

Wir betrachten nunmehr die Gesamtheit derjenigen nichtnegativen armonischen Polynome

$$P_n(r,t) = a_0 + \sum_{\nu=1}^{n} r^\nu (a_\nu \cos\nu t + b_\nu \sin\nu t) > 0 \quad \text{für} \quad r < 1,$$

ir welche $P_n(\mathsf{P}, 0) = 1$ ist ($\mathsf{P} < 1$), und fragen nach den extremen Werten on $P_n(\varrho, \varphi)$ für eine beliebige, aber festgewählte innere Stelle (ϱ, φ) des inheitskreises. Zu diesem Zwecke haben wir die extremen Werte der ermiteschen Form

38) $$P_n(\varrho,\varphi) = \sum_{\mu=0}^{n}\sum_{\nu=0}^{n} \varrho^{|\mu-\nu|} e^{(\mu-\nu)i\varphi} x_\mu \bar{x}_\nu$$

it der Nebenbedingung

39) $$P_n(\mathsf{P}, 0) = \sum_{\mu=0}^{n}\sum_{\nu=0}^{n} \mathsf{P}^{|\mu-\nu|} x_\mu \bar{x}_\nu = 1$$

u bestimmen.

Das entsprechende charakteristische Gleichungssystem ist sonach

$$\begin{aligned}
L_0(x) &\equiv (1-\lambda)x_0 + (\varrho e^{i\varphi} - \mathsf{P}\lambda)x_1 + \dots + (\varrho^n e^{ni\varphi} - \mathsf{P}^n\lambda)x_n = 0,\\
L_1(x) &\equiv (\varrho e^{-i\varphi} - \mathsf{P}\lambda)x_0 + (1-\lambda)x_1 + \dots + (\varrho^{n-1} e^{(n-1)i\varphi} - \mathsf{P}^{n-1}\lambda)x_n = 0,\\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\\
L_n(x) &\equiv (\varrho^n e^{-ni\varphi} - \mathsf{P}^n\lambda)x_0 + (\varrho^{n-1} e^{-(n-1)i\varphi} - \mathsf{P}^{n-1}\lambda)x_1 + \dots + (1-\lambda)x_n = 0.
\end{aligned}$$

Wir führen zuerst durch die homogene lineare Substitution[13])

41)
$$\begin{aligned}
X_0 &= x_0 + \mathsf{P}x_1 + \mathsf{P}^2 x_2 + \dots + \mathsf{P}^n x_n,\\
X_1 &= \mathsf{P}x_0 + x_1 + \mathsf{P}x_2 + \dots + \mathsf{P}^{n-1} x_n,\\
&\dots\dots\dots\dots\dots\dots\dots\\
X_n &= \mathsf{P}^n x_0 + \mathsf{P}^{n-1} x_1 + \mathsf{P}^{n-2} x_2 + \dots + x_n
\end{aligned}$$

---

[12]) Die Abschätzung der anderen Amplituden läßt sich bekanntlich auf die erste rückführen (vgl. L. Fejér, loc. cit. [1]) b)).

[13]) Die Substitutionsdeterminante ist $(1-\mathsf{P}^2)^{n-1}$, also für $\mathsf{P} < 1$ von 0 verschieden.

die neuen Variabeln $X_0, X_1, \ldots, X_n$ ein. Die inverse Substitution ist

$$(42)\qquad \begin{gathered} x_0 = \frac{X_0 - \mathsf{P} X_1}{1-\mathsf{P}^2}, \qquad x_n = \frac{X_n - \mathsf{P} X_{n-1}}{1-\mathsf{P}^2}, \\ x_k = \frac{(1-\mathsf{P}^2) X_k - \mathsf{P}(X_{k-1} + X_{k+1})}{1-\mathsf{P}^2} \qquad (k = 1, 2, \ldots, n-1). \end{gathered}$$

Nunmehr ersetzen wir die so entstehenden linearen Gleichungen $\bar{L}_\nu(X)$ durch die folgenden:

$$(43)\qquad \begin{gathered} -\frac{1}{\varrho}\bar{L}_0(X) + \bar{L}_1(X) = 0, \qquad \bar{L}_{n-1}(X) - \frac{1}{\varrho}\bar{L}_n(X) = 0, \\ \bar{L}_{k-1}(X) - \left(\varrho + \frac{1}{\varrho}\right)\bar{L}_k(X) + \bar{L}_{k+1}(X) = 0 \qquad (k = 1, 2, \ldots, n-1). \end{gathered}$$

Die entsprechende charakteristische Gleichung ist

$$(44)\qquad D_{n+1}(\lambda; \mathsf{P}, \varrho e^{i\varphi})$$

$$\doteqdot \begin{vmatrix} \frac{1}{\mathsf{P}} - \frac{1}{\varrho}\Lambda, & \Lambda - e^{i\varphi}, & 0 & \ldots & 0 \\ \Lambda - e^{-i\varphi}, & \frac{1}{\mathsf{P}} + \mathsf{P} - \left(\frac{1}{\varrho} + \varrho\right)\Lambda, & \Lambda - e^{i\varphi} & \ldots & 0 \\ 0, & \Lambda - e^{-i\varphi}, & \frac{1}{\mathsf{P}} + \mathsf{P} - \left(\frac{1}{\varrho} + \varrho\right)\Lambda & \ldots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0, & 0, & 0, & \ldots & \frac{1}{\mathsf{P}} - \frac{1}{\varrho}\Lambda \end{vmatrix}$$

$$\doteqdot \frac{\sin(n+2)\theta}{\sin\theta} - 2\frac{\mathsf{P} - \varrho\Lambda}{|e^{i\varphi} - \Lambda|}\frac{\sin(n+1)\theta}{\sin\theta} + \frac{(\mathsf{P} - \varrho\Lambda)^2}{|e^{i\varphi} - \Lambda|^2}\frac{\sin n\theta}{\sin\theta}$$

$$(45)\qquad \doteqdot \frac{1}{\sin\theta}\left[\cos(n+2)\frac{\theta}{2} - \frac{\mathsf{P} - \varrho\Lambda}{|e^{i\varphi} - \Lambda|}\cos n\frac{\theta}{2}\right]\left[\sin(n+2)\frac{\theta}{2} - \frac{\mathsf{P} - \varrho\Lambda}{|e^{i\varphi} - \Lambda|}\sin n\frac{\theta}{2}\right]$$

wo Λ und θ durch die Gleichungen

$$(46)\qquad \Lambda = \frac{\frac{1}{\mathsf{P}} - \mathsf{P}}{\frac{1}{\varrho} - \varrho}\lambda; \quad 2\cos\theta = \frac{\frac{1}{\mathsf{P}} + \mathsf{P} - \left(\frac{1}{\varrho} + \varrho\right)\Lambda}{|e^{i\varphi} - \Lambda|}$$

bestimmt sind.

Wir wollen nur den Fall $\mathsf{P} = \varrho$ näher betrachten. Dann ist nach (46)

$$(46')\qquad \Lambda = \lambda; \quad 2\cos\theta = \left(\frac{1}{\varrho} + \varrho\right)\frac{1-\lambda}{|e^{i\varphi} - \lambda|};$$

also

$$(47)\qquad \lambda = \frac{\left[\sin\frac{\varphi}{2}\cos\theta \pm \sqrt{\frac{1}{4}\left(\frac{1}{\varrho} + \varrho\right)^2 - \cos^2\frac{\varphi}{2}\cos^2\theta}\right]^2}{\frac{1}{4}\left(\frac{1}{\varrho} + \varrho\right)^2 - \cos^2\theta}$$

e Gleichung (44) ist hiernach für $P = \varrho$, wie auch sonst leicht ersicht-ch, eine reziproke Gleichung.

Weiterhin ergibt sich aus (45) mit Rücksicht auf (46′) durch eine nfache Rechnung

$$D_{n+1}(\lambda; \varrho, \varrho e^{i\varphi})$$

$$= \frac{1}{\sin\theta}\left[\cos(n+2)\frac{\theta}{2} - \varrho^2\cos(n-2)\frac{\theta}{2}\right]\left[\sin(n+2)\frac{\theta}{2} - \varrho^2\sin(n-2)\frac{\theta}{2}\right] = 0.$$

Man erkennt auf dieselbe Weise wie bei der Gleichung (24), daß die leichung (45′) lauter reelle, von 0 und π verschiedene Wurzeln besitzt, nd die größte sowie die kleinste Wurzel λ_M, λ_m entsprechen gemäß (47) erjenigen Wurzel θ, die zu 0 am nächsten liegt. Der zweite Faktor echts verschwindet (bei der Annahme $0 < \varrho < 1$) genau einmal, der dritte berhaupt nicht ($\theta = 0$ kommt wegen des Nenners nicht in Betracht) n Intervall $\left(0, \frac{\pi}{n+2}\right)$. Folglich ist für die in Betracht kommende Wurzel $< \theta < \frac{\pi}{n+2}$ und

48) $$\cos(n+2)\frac{\theta}{2} - \varrho^2\cos(n-2)\frac{\theta}{2} = 0.$$

Man kann also die Größe ϱ sowie die entsprechende größte und kleinste Wurzel $\lambda_M(\varphi)$ und $\lambda_m(\varphi)$ der Gleichung (45) als Funktionen des Parameters θ ausdrücken. Es ist nach (48)

49) $$\varrho = \sqrt{\frac{\cos(n+2)\frac{\theta}{2}}{\cos(n-2)\frac{\theta}{2}}} \qquad \left(0 \leqq \theta \leqq \frac{\pi}{n+2}\right),$$

nd durch Einsetzen dieses Wertes in (47)

$$= \frac{\left[\sqrt{\cos(n+2)\frac{\theta}{2}\cos(n-2)\frac{\theta}{2}}\sin\frac{\varphi}{2} + \sqrt{\cos^2 n\frac{\theta}{2} - \cos(n+2)\frac{\theta}{2}\cos(n-2)\frac{\theta}{2}\cos^2\frac{\varphi}{2}}\right]^2}{\sin^2\theta} = \frac{1}{\lambda_m(\varphi)}.$$

Es gilt demnach für ein beliebiges nichtnegatives harmonisches Polynom ter Ordnung $P_n(r, t)$

$$\frac{1}{\lambda_M(\varphi-\psi)} \leqq \frac{P_n\left(\sqrt{\frac{\cos(n+2)\frac{\theta}{2}}{\cos(n-2)\frac{\theta}{2}}}, \varphi\right)}{P_n\left(\sqrt{\frac{\cos(n+2)\frac{\theta}{2}}{\cos(n-2)\frac{\theta}{2}}}, \psi\right)} \leqq \lambda_M(\varphi-\psi) \qquad \begin{pmatrix} 0 \leqq \varphi - \psi \leqq \pi \\ 0 \leqq \theta \leqq \frac{\pi}{n+2} \end{pmatrix}.$$

Da aber einerseits eine im abgeschlossenen Kreisgebiete $\varrho \leqq \sqrt{\frac{\cos(n+2)\frac{\theta}{2}}{\cos(n-2)\frac{\theta}{2}}}$ reguläre, nichtverschwindende, harmonische Funktion ihre extremen Werte am Rande erreicht, andererseits $\lambda_M(\varphi)$ seinen größten Wert (bei gegebenen θ) offenbar für $\varphi = \pi$ annimmt, so folgt hieraus der

Satz II. *Ist das harmonische Polynom $P_n(r, t)$ im Einheitskreise $r \leqq 1$ nichtnegativ, so ist für*

$$0 \leqq r_1, r_2 \leqq \sqrt{\frac{\cos(n+2)\frac{\theta}{2}}{\cos(n-2)\frac{\theta}{2}}}, \quad 0 \leqq \theta \leqq \frac{\pi}{n+2}$$

$$(52) \quad \frac{1}{\lambda_M(\pi)} \leqq \frac{P_n(r_1, t_1)}{P_n(r_2, t_2)} \leqq \lambda_M(\pi), \qquad \lambda_M(\pi) = \left[\frac{\cos n\frac{\theta}{2} + \sqrt{\cos(n+2)\frac{\theta}{2}\cos(n-2)\frac{\theta}{2}}}{\sin\theta}\right.$$

Für das (im wesentlichen) einzige nichtnegative, harmonische Polynom $P_n^*(r, t)$, für welches in (52) Gleichheit gilt, ergibt eine unschwere Rechnung:

$$P_n^*(1, t) = \left[\frac{\sin n\frac{t+\theta}{2}}{\sin\frac{t+\theta}{2}}\cos\frac{t-\theta}{2} + \frac{\sin n\frac{t-\theta}{2}}{\sin\frac{t-\theta}{2}}\cos\frac{t+\theta}{2}\right]^2.$$

§ 5.

Wir betrachten die Gesamtheit derjenigen nichtnegativen, harmonischen Polynome n-ter Ordnung $P_n(r, t)$, welche an den Stellen $(1, \varphi)$ bzw. $(1, \psi)$ die vorgeschriebenen Werte $P_\varphi > 0$ bzw. $P_\psi > 0$ annehmen, und fragen nach dem Minimum von $P_n(\varrho, \tau)$ für eine beliebige, aber festgewählte innere Stelle (ϱ, τ) des Einheitskreises. Man kann offenbar $\tau = 0$ annehmen; man hat also das Minimum der Hermiteschen Form

$$(54) \qquad P_n(\varrho, 0) = \sum_{\mu=0}^{n}\sum_{\nu=0}^{n}\varrho^{|\mu-\nu|}x_\mu\bar{x}_\nu$$

mit den Nebenbedingungen

$$(55) \quad P_n(1, \varphi) = \sum_{\mu=0}^{n}\sum_{\nu=0}^{n}e^{(\mu-\nu)i\varphi}x_\mu\bar{x}_\nu = P_\varphi;\ P_n(1, \psi) = \sum_{\mu=0}^{n}\sum_{\nu=0}^{n}e^{(\mu-\nu)i\psi}x_\mu\bar{x}_\nu = P$$

zu bestimmen.

Führen wir in den Bedingungsgleichungen des Extremums

$$(56) \qquad \sum_{\mu=0}^{n}\{\varrho^{|\mu-\nu|} - \lambda e^{(\mu-\nu)\varphi i} - \varkappa\cdot e^{(\mu-\nu)\psi i}\}x_\mu = 0 \qquad (\nu = 0, 1, \ldots, n)$$

lie durch die Formeln (41), § 4 definierten neuen Variablen X_ν ein, so rgibt sich

$$X_\nu = \lambda \frac{1-2\varrho\cos\varphi+\varrho^2}{1-\varrho^2}\sum_{\mu=0}^{n} X_\mu e^{\mu\varphi i}\cdot e^{-\nu\varphi i} + \varkappa \frac{1-2\varrho\cos\psi+\varrho^2}{1-\varrho^2}\sum_{\mu=0}^{n} X_\mu e^{\mu\psi i}\cdot e^{-\nu\psi i}$$

$$+ \varrho\lambda\left[\frac{e^{\varphi i}-\varrho}{1-\varrho^2}X_0 + \frac{e^{-\varphi i}-\varrho}{1-\varrho^2}e^{n\varphi i}X_n\right]e^{-\nu\varphi i} + \varrho\varkappa\left[\frac{e^{\psi i}-\varrho}{1-\varrho^2}X_0 + \frac{e^{-\psi i}-\varrho}{1-\varrho^2}e^{n\psi i}X_n\right]e^{-\nu\psi i},$$

voraus man erkennt, daß einerseits die Determinante des Gleichungssystems 56) eine bilineare Form der Größen $\lambda, \varkappa$ ist, andererseits die Unbekannten X_ν lie folgende Form haben

58) $$X_\nu = A\cdot e^{-\nu\varphi i} + B\cdot e^{-\nu\psi i},$$

vo A und B von ν unabhängige Größen sind. Hieraus ergibt sich

$$\sum_{\nu=0}^{n} X_\nu e^{\nu\varphi i} = (n+1)A + e^{\frac{n}{2}(\varphi-\psi)i}\frac{\sin(n+1)\frac{\varphi-\psi}{2}}{\sin\frac{\varphi-\psi}{2}}B,$$

$$\sum_{\nu=0}^{n} X_\nu e^{\nu\psi i} = (n+1)B + e^{-\frac{n}{2}(\varphi-\psi)i}\frac{\sin(n+1)\frac{\varphi-\psi}{2}}{\sin\frac{\varphi-\psi}{2}}A$$

nd nach Einsetzen dieser Ausdrücke in (57) die Bestimmungsgleichungen ür A und B

59)
$$A[1-\lambda S_n(0,\varphi)] - B\lambda e^{n\frac{\varphi-\psi}{2}i}S_n\left(\frac{\varphi-\psi}{2},\varphi\right) = 0,$$
$$-A\varkappa e^{n\frac{\psi-\varphi}{2}i}S_n\left(\frac{\psi-\varphi}{2},\psi\right) + B[1-\varkappa S_n(0,\psi)] = 0.$$

Hier bedeutet

$$S_n(\alpha_1,\alpha_2) = \frac{(1-2\varrho\cos\alpha_2+\varrho^2)\frac{\sin(n+1)\alpha_1}{\sin\alpha_1} + 2\varrho[\cos n\alpha_1(\cos\alpha_2-\varrho) - \sin n\alpha_1\cdot\sin\alpha_2]}{1-\varrho^2}$$

$$= \frac{\sin(n+1)\alpha_1 - 2\varrho\sin n\alpha_1\cos(\alpha_1-\alpha_2)+\varrho^2\sin(n-1)\alpha_1}{(1-\varrho^2)\sin\alpha_1}.$$

s ist

$$S_n(0,\varphi) = n\frac{1-2\varrho\cos\varphi+\varrho^2}{1-\varrho^2}+1; \qquad S_n\left(\frac{\varphi-\psi}{2},\varphi\right) = S_n\left(\frac{\psi-\varphi}{2},\psi\right).$$

amit diese Gleichungen ein nichtverschwindendes Lösungssystem A, B esitzen, muß die Determinante

60)
$$\begin{vmatrix} 1-\lambda S_n(0,\varphi); & -\lambda S_n\left(\frac{\varphi-\psi}{2},\varphi\right) \\ -\varkappa S_n\left(\frac{\psi-\varphi}{2},\psi\right); & 1-\varkappa S_n(0,\psi)\end{vmatrix}$$

die obengenannte Bilinearform von $\lambda, \varkappa$) verschwinden.

Dann ist

$$(61)\qquad \frac{Be^{-n\frac{\psi}{2}i}}{Ae^{-n\frac{\varphi}{2}i}} = \frac{1-\lambda S_n(0,\varphi)}{\lambda S_n\left(\frac{\varphi-\psi}{2},\varphi\right)} = \frac{\varkappa S_n\left(\frac{\psi-\varphi}{2},\psi\right)}{1-\varkappa S_n(0,\psi)}.$$

Sind die Gleichungen (59), also auch die Gleichungen (56) befriedigt, so ist

$$P_n(\varrho,0) = \lambda P_\varphi + \varkappa P_\psi,$$

und man hat noch das Minimum der Linearform $\lambda P_\varphi + \varkappa P_\psi$ mit der Nebenbedingung (60) zu bestimmen. Hierfür ergibt sich durch eine elementare Rechnung

$$(62)\quad \operatorname{Min} P_n(\varrho,0) = \frac{P_\varphi S_n(0,\psi) + P_\psi S_n(0,\varphi) - 2\sqrt{P_\varphi P_\psi}\left|S_n\left(\frac{\varphi-\psi}{2},\varphi\right)\right|}{S_n(0,\varphi)S_n(0,\psi) - \left[S_n\left(\frac{\varphi-\psi}{2},\varphi\right)\right]^2}.$$

Durch Einsetzen derjenigen Werte von $\lambda, \varkappa$, für welche $\lambda P_\varphi + \varkappa P_\psi$ Minimum wird, in (61), erhält man weiterhin

$$\frac{Ae^{-\frac{n}{2}\varphi i}}{Be^{-\frac{n}{2}\psi i}} = \frac{\varkappa}{\lambda}\sqrt{\frac{P_\psi}{P_\varphi}} = \frac{\sqrt{P_\psi}S_n(0,\varphi) - \sqrt{P_\varphi}S_n\left(\frac{\varphi-\psi}{2},\varphi\right)}{\sqrt{P_\varphi}S_n(0,\psi) - \sqrt{P_\psi}S_n\left(\frac{\psi-\varphi}{2},\psi\right)},$$

also

$$(63)\quad X_\nu = \left[\sqrt{P_\psi}S_n(0,\varphi) - \sqrt{P_\varphi}S_n\left(\frac{\varphi-\psi}{2},\varphi\right)\right]e^{\left(\frac{n}{2}-\nu\right)\varphi i} + \left[\sqrt{P_\varphi}S_n(0,\psi) - \sqrt{P_\psi}S_n\left(\frac{\psi-\varphi}{2},\psi\right)\right]e^{\left(\frac{n}{2}-\nu\right)\psi i},\quad (\nu = 0,1,\ldots,n).$$

Das einzige nichtnegative harmonische Polynom $P_n^*(r,t)$, für welches das Minimum erreicht wird, bestimmt sich (nachdem man mit Berücksichtigung von (63) und der inversen Formeln (42) des § 4 die Unbekannten x_ν berechnet hat) aus der Gleichung

$$(64)\qquad P_n^*(1,t) = \left\{\frac{\left[\sqrt{P_\psi}S_n(0,\varphi) - \sqrt{P_\varphi}S_n\left(\frac{\varphi-\psi}{2},\varphi\right)\right]S_n\left(\frac{\psi-t}{2},\psi\right)}{S_n(0,\varphi)S_n(0,\psi) - S_n\left(\frac{\varphi-\psi}{2},\varphi\right)S_n\left(\frac{\psi-\varphi}{2},\psi\right)} + \frac{\left[\sqrt{P_\varphi}S_n(0,\psi) - \sqrt{P_\psi}S_n\left(\frac{\psi-\varphi}{2},\psi\right)\right]S_n\left(\frac{\varphi-t}{2},\varphi\right)}{S_n(0,\varphi)S_n(0,\psi) - S_n\left(\frac{\varphi-\psi}{2},\varphi\right)S_n\left(\frac{\psi-\varphi}{2},\psi\right)}\right\}^2,$$

welche auch in der Form (16) geschrieben werden kann.

Wird in (62) und (64) $\varrho = 0$ gesetzt, so ergeben sich die für diesen Fall von Herrn S. Bernstein angegebenen Ausdrücke[14]).

Läßt man in der Gleichung (60), nachdem man dort $\lambda = \frac{1}{\Lambda(\varphi-\psi)}$, $\varkappa = \frac{1}{\Lambda(\psi-\varphi)}$ gesetzt hat, ψ nach φ konvergieren, so erhält man

$$(65)\quad \Lambda_m . P(\varrho, 0) \leqq P_t'(1, \varphi) = \lim_{\psi \to \varphi} \frac{P(1, \varphi) - P(1, \psi)}{\varphi - \psi} \leqq \Lambda_M . P(\varrho, 0),$$

wobei Λ_m bzw. Λ_M die kleinere bzw. größere Wurzel der quadratischen Gleichung

$$\Lambda^2 - \Lambda \lim_{\psi \to \varphi} \frac{S_n(0, \varphi) - S_n(0, \psi)}{\varphi - \psi} - \lim_{\psi \to \varphi} \frac{S_n(0, \varphi) S_n(0, \psi) - \left[S_n\left(\frac{\varphi-\psi}{2}, \varphi\right)\right]^2}{(\varphi-\psi)^2} = 0$$

bedeutet.

Eine kurze Rechnung ergibt für Λ_m bzw. Λ_M

$$\Lambda = \frac{n\varrho \sin\varphi}{1-\varrho^2}$$

$$\sqrt{\frac{1}{12}[(n+1) - 2n\varrho\cos\varphi + (n-1)\varrho^2][n(n+1)(n+2) - 2(n-2)n(n+2)\varrho\cos\varphi + (n-2)(n-1)n\varrho^2]}$$

$$1-\varrho^2$$

Für $\varrho = 0$ erhält man hieraus

$$-\sqrt{\frac{n(n+1)^2(n+2)}{12}} P(0, 0) \leqq P_t'(1, \varphi) \leqq \sqrt{\frac{n(n+1)^2(n+2)}{12}} P(0, 0),$$

also die in einer gemeinsamen Arbeit von Herrn O. Szász und vom Verfasser enthaltene Abschätzung[15]) für die Derivierte eines nichtnegativen trigonometrischen Polynoms.

---

[14]) Die Bernsteinsche Formel (l. c. [5]), Formel (39)) für das Minimum von $P_n(0, 0)$ lautet nämlich, mit den hier benutzten Bezeichnungen

$$\min P_n(0, 0) = \frac{(n+1)(P_\varphi + P_\psi) - 2\left|\frac{\sin(n+1)\frac{\varphi-\psi}{2}}{\sin\frac{\varphi-\psi}{2}}\right| \sqrt{P_\varphi P_\psi}}{(n+1)^2 - \left(\frac{\sin(n+1)\frac{\varphi-\psi}{2}}{\sin\frac{\varphi-\psi}{2}}\right)^2}.$$

[15]) loc. cit. S. 650—652.

(Eingegangen am 1. Juli 1931.)

Über die Abbildungen, die durch Systeme von analytischen Funktionen von mehreren Veränderlichen erzeugt werden.

Von

C. Carathéodory in München*).

Inhalt.

Einleitung, 1.

Kapitel I. Einleitende Betrachtungen und Hilfssätze.
Vorbemerkungen, 2, 3.
Eigenschaften der Abbildung in der Umgebung eines Punktes, 4, 5.
Konvergente Abbildungsfolgen in der Umgebung eines Punktes, 6, 7.
Das Schwarzsche Lemma, 8.
Die Grenzfunktion einer Folge nicht-verschwindender Funktionen, 9.

Kapitel II. Der Kern einer Gebietsfolge.
Ein Abbildungssatz, 10—12.
Der Unitätssatz, 13.
Definition des Kerns, 14—16.
Der Hauptsatz, 17—19.

Kapitel III. Die analytischen Selbstabbildungen eines Gebietes.
Ein Konvergenzsatz, 20.
Die Menge der iterierten Selbsttransformationen, 21.
Ein Satz von Henri Cartan, 22—25.
Über die konforme Abbildung mehrfach zusammenhängender ebener Gebiete, 26, 27.

Kapitel IV. Die Maximalteiler eines Gebietes.
Definition, 28.
Eigenschaften der Maximalteiler, 29—31.
Hyperkugel und Polyzylinder, 32, 33.

Einleitung.

1. Neuerdings ist das Studium des Problems der Abbildungen von mehrdimensionalen Räumen durch Systeme von analytischen Funktionen

*) Der Inhalt dieser Arbeit bildete den Gegenstand eines am 25. Juli 1931 an der Universität Leipzig gehaltenen Vortrages.

n komplexen Gebiet mit Hilfe von sehr originellen und sinnreichen Methoden in Angriff genommen worden[1]).

Dagegen ist der Versuch noch nicht gemacht worden, die allgemeinen ätze, die bei der modernen Theorie der konformen Abbildung in der bene eine hervorragende Rolle spielen, systematisch auf die analytischen bbildungen mehrdimensionaler Gebiete zu übertragen. Dies soll auf den lgenden Seiten nachgeholt werden.

Ein Teil gerade der wichtigsten unter diesen Sätzen beruht einzig und llein auf der stetigen Konvergenz von Folgen von Abbildungsfunktionen, nd es ist daher fast selbstverständlich, daß eine vorsichtige Modifizierung on bekannten Beweismethoden zur gewünschten Verallgemeinerung führen uß. Man kann auf diese Weise die hauptsächlichsten Resultate, die ich or zwanzig Jahren in einer Arbeit über konforme Abbildungen bewiesen atte[2]), auf die allgemeinen analytischen Abbildungen übertragen. Diese ätze sind sehr bequem, und die neuen Beweise sogar einfacher als die lten. Unter den Anwendungen, die von der allgemeinen Theorie gemacht erden, befindet sich ein sehr interessanter Satz des Herrn Henri Cartan, en dieser mit speziellen Methoden für den vierdimensionalen Raum erhalten hat, und der ohne besondere Mühe auf den $2n$-dimensionalen Raum bertragen werden kann. Es ist bemerkenswert, daß derart komplizierte Sätze it den hier verwendeten elementaren Mitteln bewiesen werden können.

Der letzte Abschnitt über Hyperkugel und Polyzylinder kann auch für ich gelesen werden. Er enthält den elementarsten mir bekannten Beweis, aß diese beiden Gebiete nicht aufeinander analytisch abbildbar sind.

Kapitel I.

Einleitende Betrachtungen und Hilfssätze.

2. Vorbemerkungen. Die Gebiete, die wir im folgenden betrachten erden, sollen im $2n$-dimensionalen Raume liegen, der durch die reellen nd imaginären Teile von n komplexen Veränderlichen $(x_1, \ldots, x_n)$ erzeugt wird.

Sind nun n Funktionen

(2.1) $$f_i(x_1, \ldots, x_n) \qquad (i = 1, \ldots, n)$$

[1]) Für die Literatur bis 1930 siehe H. Behnke, Analytische Abbildungen im aume zweier komplexer Veränderlichen. Die Entwicklung der Theorie in der letzten eit, Jahresber. der D. M.-V. 40 (1931), S. 15—21. Seitdem sind außer den weiter nten zitierten Arbeiten von H. Cartan noch folgende hinzugekommen: P. Thullen, ie Starrheit der nicht überall pseudokonvexen Gebiete. Math. Annalen 104 (1931). . Siewert, Die Abbildungen der Hartogschen Körper, Diss., Münster i. W. (1930).

[2]) C. Carathéodory, Untersuchungen über die konformen Abbildungen von festen nd veränderlichen Gebieten, Math. Annalen 72 (1912), S. 107—144.

gegeben, die in einem derartigen Gebiete G des Raumes der $(x_1, \ldots, x_n)$ analytisch sind, so stellen die Gleichungen

$$x_i' = f_i(x_1, \ldots, x_n) \qquad (i = 1, \ldots, n) \tag{2.2}$$

eine analytische Abbildung auf eine Punktmenge G' des Raumes der $(x_1', \ldots, x_n')$ dar.

Es ist oft bequem, die Abbildung, die durch die Gleichungen (2.2) definiert wird, durch einen einzigen Buchstaben, wie z. B. durch S, zu bezeichnen. Um dann die Tatsache auszudrücken, daß ein Punkt P mit den Koordinaten $(x_1, \ldots, x_n)$ durch (2.2) in einen Punkt P' mit den Koordinaten $(x_1', \ldots, x_n')$ übergeht, schreibt man einfach

$$P' = S\,P. \tag{2.3}$$

Ähnlich kann man durch

$$A' = S\,A \tag{2.4}$$

ausdrücken, daß eine beliebige Teilmenge A von G durch die Transformation S in die Punktmenge A' übergeführt wird.

Wir werden nun hauptsächlich Folgen

$$S_1, S_2, \ldots \tag{2.5}$$

von analytischen Abbildungen zu betrachten haben, die mit Hilfe von Gleichungen der Form

$$x_i' = f_i^{(k)}(x_1, \ldots, x_n) \qquad (i = 1, \ldots, n) \tag{2.6}$$

dargestellt werden, und durch welche ein festes Gebiet G in die Punktmengen

$$G_k' = S_k\,G \qquad (k = 1, 2, \ldots) \tag{2.7}$$

transformiert wird.

Die Grundlage aller unserer Untersuchungen beruht nun darauf, daß für alle Folgen von Abbildungen (2.5), die wir zu betrachten haben werden, die vorkommenden Transformationsfunktionen $f_i^{(k)}(x_1, \ldots, x_n)$ einer normalen Familie angehören sollen.

Dies ist z. B. gewährleistet, wenn man von vornherein weiß, daß alle Punktmengen G_k', die in (2.7) vorkommen, Teilmengen eines festen beschränkten Gebietes H' sind. Es ist auch nützlich, von G selbst vorauszusetzen, daß es beschränkt ist, da es sonst vorkommen könnte, daß sämtliche analytischen beschränkten Abbildungen von G immer eine identisch verschwindende Funktionaldeterminante besitzen.

Wir werden also im folgenden immer voraussetzen, daß alle vorkommenden Gebiete beschränkt und schlicht sind. Viele unserer Resultate

behalten aber ihre Gültigkeit unter weitergehenden Bedingungen[3]) insbesondere auch, wenn Teile des Raumes durch die Gebiete G und H mehrfach überdeckt werden, oder, wenn die Gebiete verzweigt sind. Dies gehört aber zu einem anderen Kapitel der Funktionentheorie, das ich hier nicht anschneiden möchte.

Bei unseren Annahmen ist es immer möglich, aus der Folge der Transformationen (2.5) Teilfolgen $S_{k_1}, S_{k_2}, \ldots$ auszusondern, die im Gebiete G gegen eine analytische Abbildung S stetig konvergieren. Dies ist unter den gemachten Voraussetzungen immer der Fall, wenn die Relationen

$$\lim_{m=\infty} f_i^{(k_m)}(x_1, \ldots, x_n) = f_i(x_1, \ldots, x_n) \tag{2.8}$$

gleichzeitig für $(i = 1, \ldots, n)$ und für jeden Punkt von G stattfinden.

Ist dann irgendeine Folge von Punkten $P_1, P_2, \ldots$ gegeben, die gegen einen *inneren* Punkt P von G konvergiert, so hat man immer

$$\lim_{m=\infty} S_{k_m}(P_m) = SP. \tag{2.9}$$

3. Es sei A eine beliebige abgeschlossene Punktmenge, die ganz im Innern des beschränkten Gebietes G liegt. Dann gibt es eine feste Zahl δ, so daß für alle Punkte $(x_1^0, \ldots, x_n^0)$ von A der Polyzylinder

$$|x_i - x_i^0| < \delta$$

ebenfalls in G enthalten ist. Man kann mit Hilfe des Schwarzschen Lemmas (oder des Cauchyschen Integrals) eine gemeinsame obere Schranke für den absoluten Betrag der ersten partiellen Ableitungen $\frac{\partial f_i^{(k)}}{\partial x_j}$ aller Funktionen, die in (2.6) vorkommen, berechnen, die gleichmäßig für alle Punkte von A gilt.

Hieraus folgt dann, daß der absolute Betrag der Funktionaldeterminante

$$\frac{\partial(f_1^{(k)}, \ldots, f_n^{(k)})}{\partial(x_1, \ldots, x_n)} \tag{3.2}$$

für alle Funktionen unserer Familie und für alle Punkte von A gleichmäßig beschränkt ist.

Endlich kann man ähnliche Schranken auch für die höheren partiellen Ableitungen einer beliebigen aber festen Ordnung bestimmen, die gleichmäßig für alle Funktionen $f_i^{(k)}$ und alle Punkte von A gelten.

---

[3]) Die Voraussetzung, daß G und H beschränkt sein sollen, kann man ohne Schwierigkeit in den meisten Fällen durch die viel allgemeinere ersetzen, daß die Euklidische Entfernung zwischen zwei inneren Punkten mit der Distanz $D_G(P, Q)$ bzw. $D_H(P', Q')$ (siehe 8) gegen Null konvergiert.

4. Eigenschaften der Abbildung in der Umgebung eines Punktes. Ist in einem Punkte P_0 von G die Funktionaldeterminante

$$\Delta(P_0) = \frac{\partial(f_1, \ldots, f_n)}{\partial(x_1, \ldots, x_n)} \qquad (x_i = x_i^0) \tag{4.1}$$

von Null verschieden, so kann man die Gleichungen (2.2) in einer Umgebung dieses Punktes nach $(x_1, \ldots, x_n)$ auflösen. In diesem Falle ist die Abbildung

$$G' = SG \tag{4.2}$$

in der Umgebung von P_0 *topologisch*, d. h. sie ist erstens *gebietstreu*, zweitens *eineindeutig* und drittens *umkehrbar stetig*.

5. Es ist sehr wahrscheinlich, daß die zweite dieser Eigenschaften allein schon hinreicht, um das Nichtverschwinden der Funktionaldeterminante zu gewährleisten. Hierzu müßte man den folgenden Satz beweisen: *ist die Funktionaldeterminante $\Delta(P_0) = 0$, so gibt es in jeder Umgebung von P_0 Punktepaare P, Q, die dasselbe Bild besitzen.*

Der Beweis dieser letzten Tatsache ist aber jedenfalls nicht einfach[4]). Glücklicherweise können wir uns im folgenden mit einem spezielleren Satz begnügen, der implizite in längst bekannten Resultaten enthalten ist und für dessen direkte Ableitung mir Herr F. Hartogs eine sehr einfache Schlußweise mündlich mitgeteilt hat.

Es gilt der

Satz 1. *Sind die analytischen Funktionen*

$$\varphi_1(x_1, \ldots, x_n), \quad \ldots, \quad \varphi_n(x_1, \ldots, x_n)$$

regulär im Punkte $(x_1^0, \ldots, x_n^0)$ und wird durch die Gleichungen

$$u_i = \varphi_i(x_1, \ldots, x_n) \qquad (i = 1, \ldots, n) \tag{5.1}$$

eine Umgebung dieses Punktes eineindeutig und umkehrbar stetig auf eine Umgebung der Stelle $(u_1^0, \ldots, u_n^0)$ abgebildet, so besteht in $(x_1^0, \ldots, x_n^0)$ die Relation

$$\frac{\partial(\varphi_1, \ldots, \varphi_n)}{\partial(x_1, \ldots, x_n)} \neq 0. \tag{5.2}$$

Nach Voraussetzung gibt es stetige Funktionen

$$x_j = \psi_j(u_1, \ldots, u_n) \qquad (j = 1, \ldots, n), \tag{5.3}$$

die in einer Umgebung des Bildpunktes $(u_1^0, \ldots, u_n^0)$ von $(x_1^0, \ldots, x_n^0)$ definiert sind und die, in (5.1) eingesetzt, diese Gleichungen identisch befriedigen. Gelingt es nun zu zeigen, daß diese Funktionen ψ_j im erwähnten Gebiet

[4]) Für zwei Veränderliche folgt der Satz aus den Entwicklungen auf S. 6–8 der Arbeit von H. Cartan, Les fonctions de deux variables complexes et le problème de la représentation analytique, Journ. de Mathém. **10** (1931), S. 1–114.

reguläre analytische Funktionen sind, so ist die behauptete Relation (5.2) evident, weil das Produkt der Funktionaldeterminanten

$$\frac{\partial(\varphi_1, \ldots, \varphi_n)}{\partial(x_1, \ldots, x_n)} \cdot \frac{\partial(\psi_1, \ldots, \varphi_n)}{\partial(u_1, \ldots, u_n)} \equiv 1 \tag{5.4}$$

ist.

Aus den Voraussetzungen des Satzes 1 folgt aber schon von selbst, daß die Determinante (5.2) nicht in einer vollen Umgebung von $(x_1^0, \ldots, x_n^0)$ identisch verschwinden kann. In diesem Falle würde nämlich mindestens eine analytische Relation

$$\Omega(u_1, \ldots, u_n) = 0$$

zwischen den Größen u_i bestehen, und die Punkte $(u_1, \ldots, u_n)$ könnten dann nicht eine Umgebung des Punktes $(u_1^0, \ldots, u_n^0)$ ausfüllen.

Die Gleichung

$$\frac{\partial(\varphi_1, \ldots, \varphi_n)}{\partial(x_1, \ldots, x_n)} = 0$$

kann also höchstens auf einer Punktmenge $\mathfrak{C}$ erfüllt sein, die aus einem oder mehreren charakteristischen Gebilden von $(2n-2)$-ter oder niedrigerer Dimension besteht. Diese Punktmenge $\mathfrak{C}$ wird durch die analytischen Funktionen (5.1) auf eine Punktmenge $\mathfrak{C}^*$ abgebildet, die die gleichen Eigenschaften hat. In jedem Punkte einer gewissen Umgebung von $(u_1^0, \ldots, u_n^0)$, außer vielleicht in den Punkten von $\mathfrak{C}^*$ selbst, sind die beschränkten und stetigen Funktionen (5.3) reguläre analytische Funktionen der $(u_1, \ldots, u_n)$. Also müssen sie nach dem verallgemeinerten Riemannschen Satz[5]) auch in den Punkten von $\mathfrak{C}^*$, insbesondere im Punkte $(u_1^0, \ldots, u_n^0)$ analytisch sein, womit die Gleichung (5.4) und infolgedessen auch die Relation (5.2) bewiesen ist.

6. Konvergente Abbildungsfolgen in der Umgebung eines Punktes. Wir betrachten, ähnlich wie im § 2, eine stetig konvergente Folge von Transformationen

$$S_1, S_2, \ldots, \tag{6.1}$$

die alle in einer Umgebung G eines Punktes P_0 definiert sind, und nehmen an, daß der Wert $\Delta(P_0)$ der Funktionaldeterminante der Grenztransformation S im Punkte P_0 von Null verschieden ist.

Unter dieser Voraussetzung gibt es eine abgeschlossene Hyperkugel $\varkappa^*$ mit dem Mittelpunkte P_0 und dem Radius ϱ^*, die durch die Transformation S *topologisch*, d. h. eineindeutig und umkehrbar stetig, auf ein abgeschlossenes Gebiet $S\varkappa^*$ abgebildet wird. Wenn ein Punkt P die abgeschlossene Punktmenge $\varkappa^*$ beschreibt, so wird der absolute Betrag der

[5]) Siehe z. B. W. F. Osgood, Lehrbuch der Funktionentheorie 2, zweite Auflage, Leipzig 1929, S. 186.

Funktionaldeterminante $\Delta(P)$ von S eine von Null verschiedene untere Grenze haben.

Da nun die Folge (6.1) der Transformationen S_k in $\varkappa^*$ stetig konvergiert, gibt es eine positive Zahl k^*, so daß für alle $k > k^*$ und alle Punkte P aus $\varkappa^*$, die Funktionaldeterminante $\Delta_k(P)$ von S_k von Null verschieden ist.

Man kann aber leicht ein noch viel genaueres Resultat ableiten: wir wollen zeigen, daß es eine in $\varkappa^*$ enthaltene und zu dieser konzentrische Hyperkugel $\varkappa_0$ gibt, sowie eine Zahl $k_0 \geqq k^*$, so daß, für alle $k \geqq k_0$, jede der Transformationen S_k eine *topologische* Abbildung von $\varkappa_0$ liefert.

Im entgegengesetzten Falle würde es nämlich eine monotone Folge von natürlichen Zahlen $k_1, k_2, \ldots$ geben, sowie zwei gegen P_0 konvergierende Folgen von Punkten $Q_1, Q_2, \ldots$ und $R_1, R_2, \ldots$, so daß die beiden Punkte Q_m und R_m voneinander verschieden sind und doch durch S_{k_m} in *denselben* Punkt transformiert werden.

Bezeichnet man mit $\xi_i^{(m)}$ und $\eta_i^{(m)}$ die Koordinaten von Q_m bzw. von R_m und mit y_i die Koordinaten von P_0, so würden dann erstens die Relationen

$$\lim_{m=\infty} \xi_i^{(m)} = \lim_{m=\infty} \eta_i^{(m)} = y_i \qquad (i = 1, \ldots, n), \tag{6.2}$$

$$|\eta_1^{(m)} - \xi_1^{(m)}|^2 + \ldots + |\eta_n^{(m)} - \xi_n^{(m)}|^2 = r_m^2, \tag{6.3}$$

$$r_m \neq 0, \qquad \lim_{m=\infty} r_m = 0 \tag{6.4}$$

bestehen, und zweitens würden, mit den Bezeichnungen von 2., die Gleichungen

$$f_i^{(k_m)}(\eta_1^{(m)}, \ldots, \eta_n^{(m)}) - f_i^{(k_m)}(\xi_1^{(m)}, \ldots, \xi_n^{(m)}) = 0 \tag{6.5}$$

verifiziert sein.

Nun setze man

$$\alpha_j^{(m)} = \frac{\eta_j^{(m)} - \xi_j^{(m)}}{r_m} \qquad (j = 1, \ldots, n) \tag{6.6}$$

und

$$\varphi_i^{(m)}(t) = f_i^{(k_m)}(\xi_j^{(m)} + \alpha_j^{(m)} t) - f_i^{(k_m)}(\xi_j^{(m)}). \tag{6.7}$$

Wir können immer erzwingen, indem wir eventuell die Folge der S_{k_m} durch eine Teilfolge ersetzen, daß die Grenzwerte

$$\lim_{m=\infty} \alpha_j^{(m)} = \alpha_j \tag{6.8}$$

alle existieren. Dann konvergieren in einer gewissen Umgebung des Punktes $t = 0$ die Funktionen (6.7) stetig gegen die Funktionen

$$\varphi_i(t) = f_i(y_j + \alpha_j t) - f_i(y_j). \tag{6.9}$$

Man bemerke, daß nach (6.5) bis (6.7)

$$\varphi_i^{(m)}(r_m) = \varphi_i^{(m)}(0) = 0$$

ist, und daß man daher, mit Berücksichtigung von (6.4), schreiben kann

$$\lim_{m=\infty} \frac{\varphi_i^{(m)}(r_m) - \varphi_i^{(m)}(0)}{r_m} = \varphi_i'(0) = 0.$$

Aus (6.9) folgt nunmehr

$$\alpha_1 \frac{\partial f_i}{\partial x_1} + \ldots + \alpha_n \frac{\partial f_i}{\partial x_n} = 0 \qquad (i = 1, \ldots, n) \tag{6.10}$$

im Punkte P_0. Andererseits ist nach (6.3), (6.6) und (6.8)

$$|\alpha_1^2| + \ldots + |\alpha_n^2| = 1;$$

die α_j können also nicht durchweg verschwinden, und die Determinante des Systems der linearen, homogenen Gleichungen (6.10) muß daher Null sein. Diese Determinante ist aber identisch mit unserer Funktionaldeterminante $\Delta(P_0)$, die nach Voraussetzung nicht verschwinden sollte; es liegt also ein Widerspruch vor, aus welchem die Richtigkeit unserer ursprünglichen Behauptung folgt.

7. Wir bezeichnen nun mit B_0 den Rand der Hyperkugel $\varkappa_0$, und führen die Bezeichnungen ein

$$B_0' = S B_0, \qquad B_k' = S_k B_0, \tag{7.1}$$

$$P_0' = S P_0, \qquad P_k' = S_k P_0. \tag{7.2}$$

Nach dem Resultate des letzten Paragraphen ist nun sowohl die Entfernung δ zwischen dem Punkte P_0' und der abgeschlossenen Punktmenge B_0' von Null verschieden, als auch, für $k \geqq k_0$, die Entfernung δ_k zwischen P_k' und B_k'. Außerdem hat man, wegen der stetigen Konvergenz der Abbildungsfolge $S_1, S_2, \ldots$,

$$\lim_{k=\infty} \delta_k = \delta, \qquad \lim_{k=\infty} P_k' = P_0'. \tag{7.3}$$

Endlich stellt, für $k \geqq k_0$, die Punktmenge B_k' die Begrenzung des abgeschlossenen Gebietes $S_k \varkappa_0$ dar, und ebenso ist B_0' die Begrenzung von $S\varkappa_0$.

Hieraus entnimmt man erstens, daß, für $k \geqq k_0$, die Hyperkugel mit dem Mittelpunkt P_k' und dem Radius δ_k im Gebiete $S_k \varkappa_0$ enthalten ist, und daß die Hyperkugel mit dem Mittelpunkte P_0' und dem Radius δ in $S\varkappa_0$ liegt.

Wegen (7.3) gibt es also eine natürliche Zahl N, so daß für alle $k \geqq N$ die Hyperkugel mit dem Mittelpunkte P_0' und dem Radius $\frac{\delta}{2}$ in allen Gebieten $S_k \varkappa_0$ enthalten ist, und wir können mithin den folgenden Satz aussprechen:

Satz 2. *Es sei eine stetig konvergente Folge von analytischen Transformationen $S_1, S_2, \ldots$ gegeben. In einem Punkte P_0 sei die Funktionaldeterminante der Grenztransformation S von Null verschieden. Dann kann*

man eine Umgebung U_{P_0} von P_0, eine Umgebung $U_{P_0'}$ des Bildpunktes $P_0' = S P_0$ und eine natürliche Zahl N angeben, die folgenden Bedingungen genügen: Die Abbildung des Gebietes U_{P_0} durch die Transformation S und durch alle Transformationen $S_N, S_{N+1}, \ldots$ ist topologisch, und das feste Gebiet $U_{P_0'}$ ist eine Teilmenge aller durch diese Transformationen gewonnenen Bilder von U_{P_0}.

Dieser Satz ist einer der wichtigsten, die in der Theorie der analytischen Abbildungen vorkommen: er stellt in der Tat die richtige Verallgemeinerung des Satzes über die Inversion eines Funktionensystems bei Nichtverschwinden der Funktionaldeterminante dar, wenn man an Stelle eines einzigen Funktionensystems eine stetig konvergente Folge derartiger Systeme betrachtet.

8. Das Schwarzsche Lemma. Vor einigen Jahren habe ich gezeigt, daß der Wunsch, das Schwarzsche Lemma auch auf analytische Funktionen von mehreren Veränderlichen zu übertragen, dazu führt, jedem $2n$-dimensionalen Gebiet eine ihm eigentümliche Metrik zuzuordnen[6]).

Hierzu betrachtet man die Gesamtheit der analytischen Funktionen $\varphi(x_1, \ldots, x_n)$, die im Gebiete G regulär und absolut genommen kleiner als Eins sind, und definiert als Distanz $D_G(A, B)$ von zwei Punkten A und B von G die obere Grenze der nichteuklidischen Entfernungen $E(z_A, z_B)$ der Punkte z_A und z_B des Einheitskreises, die die Werte einer jeden der Funktionen φ in den Punkten A und B darstellen.

Herr Henri Cartan hat nun bemerkt[7]), daß man mit Hilfe dieser Metrik einen Satz erhält, der der ursprünglichen Fassung des Schwarzschen Lemmas sehr ähnlich ist und der folgendermaßen ausgesprochen werden kann:

Satz 3. *Bezeichnet man mit S eine analytische Abbildung, die das Gebiet G in eine Punktmenge G' transformiert, die im Innern eines Gebietes H liegt, so besteht, wenn A und B zwei beliebige Punkte von G bedeuten und wenn man*

$$A' = S A, \qquad B' = S B$$

setzt, immer die Relation

$$D_H(A'B') \leqq D_G(AB). \tag{8.1}$$

Es gibt nämlich mindestens eine in H reguläre analytische Funktion $\psi(x_1', \ldots, x_n')$, deren absoluter Betrag in diesem Gebiete < 1 ist, derart,

---

[6]) C. Carathéodory, Über das Schwarzsche Lemma bei analytischen Funktionen von zwei komplexen Veränderlichen, Math. Annalen **97** (1927), S. 76–98.

[7]) Henri Cartan, Sur les fonctions de deux variables complexes. Les transformations d'un domaine borné D en un domaine intérieur à D, Bull. Soc. math. de France, **58** (1930), S. 199—219, insbes. S. 205.

laß die nichteuklidische Entfernung

$$E(\psi(A'), \psi(B')) = D_H(A', B') \tag{8.2}$$

st. Wird nun die Transformation S durch die Gleichungen

$$x_i' = f_i(x_1, \ldots, x_n) \tag{8.3}$$

largestellt und setzt man

$$\varphi(x_1, \ldots, x_n) = \psi(f_1, \ldots, f_n), \tag{8.4}$$

o ist die Funktion φ analytisch in G und dort absolut genommen < 1. Es ist daher nach Definition

$$E(\varphi(A), \varphi(B)) \leqq D_G(A, B). \tag{8.5}$$

Die Relation (8.1) ist nun eine Folge von (8.2) und (8.5), wenn man bedenkt, daß

$$\varphi(A) = \psi(A'), \quad \varphi(B) = \psi(B') \tag{8.5}$$

st.

In der unter [6]) zitierten Arbeit wurde gezeigt, daß man jedem rektifizierbaren abgeschlossenen Kurvenstück γ, das ganz im Innern von G enthalten ist, eine „metrische Länge"

$$L_G(\gamma)$$

uordnen kann. Aus dem letzten Satze folgt dann ohne weiteres, daß, wenn man

$$\gamma' = S\gamma \tag{8.7}$$

etzt, stets die Relation

$$L_H(\gamma') \leqq L_G(\gamma) \tag{8.8}$$

rfüllt sein muß. Aus dieser Bemerkung entnimmt man den

Satz 4. *Die analytische Abbildung S genüge den Voraussetzungen des Satzes 3; ferner sei, wenn O einen inneren Punkt von G bedeutet,*

$$O' = SO.$$

Bezeichnet man mit G_ϱ die Punkte von G, die mit O durch eine Kurve γ von metrischer Länge

$$L_G(\gamma) < \varrho$$

verbunden werden können, wobei ϱ eine beliebige positive Zahl bedeutet, und ebenso mit H_ϱ die Punkte von H, die mit O' durch eine Kurve δ' von metrischer Länge

$$L_H(\delta') < \varrho$$

verbunden werden können, so ist immer

$$G_\varrho' = S G_\varrho$$

eine Teilmenge von H_ϱ.

9. Die Grenzfunktion einer Folge nicht-verschwindender Funktionen. Der letzte Hilfssatz, den wir für die Darstellung der allgemeinen Theorie benötigen, ist folgender

Satz 5. *Es seien alle analytischen Funktionen der Folge*

$$\varphi_1(x_1, \ldots, x_n),\ \varphi_2(x_1, \ldots, x_n), \ldots$$

regulär und von Null verschieden in sämtlichen Punkten eines Gebietes G; ferner sei diese Folge überall in G stetig konvergent. Dann ist die Grenzfunktion $\varphi(x_1, \ldots, x_n)$ entweder identisch Null oder immer von Null verschieden.

Für $n = 1$ folgt dieser Satz aus einem sehr bekannten Satz von Hurwitz; man kann ihn aber auch direkt in wenigen Zeilen wie folgt beweisen: ist die Grenzfunktion $\varphi(x)$ konstant, so ist die Behauptung des Satzes richtig. Ist aber $\varphi(x)$ keine Konstante, so kann man um jeden gegebenen Punkt x_0 von G einen Kreis $|x - x_0| = r$ zeichnen, der in G liegt und auf welchem $\varphi(x) \neq 0$ ist. Es gibt dann immer eine Zahl $\varepsilon > 0$, so daß auf der gewählten Kreislinie die Relation $|\varphi(x)| \geqq \varepsilon$ erfüllt ist. Dann ist für hinreichend große Werte von k, wegen der gleichmäßigen Konvergenz, auch $|\varphi_k(x)| \geqq \frac{\varepsilon}{2}$ auf demselben Kreis. Die Funktionen $\varphi_k(x)$ sind nun nach Voraussetzung alle von Null verschieden auf der Kreisfläche $|x - x_0| \leqq r$, und man hat daher, für die soeben erwähnten Werte von k,

$$|\varphi_k(x_0)| \geqq \frac{\varepsilon}{2}.$$

Hieraus folgt sofort durch Grenzübergang die behauptete Tatsache, daß $\varphi(x_0) \neq 0$ ist.

Den Fall $n > 1$ führt man auf den soeben behandelten zurück. Man kann wieder ohne Einschränkung der Allgemeinheit voraussetzen, daß die Grenzfunktion $\varphi(x_1, \ldots, x_n)$ nicht konstant ist. In diesem Falle gibt es, wenn unter $(x_1^0, \ldots, x_n^0)$ die Koordinaten eines beliebigen Punktes von G verstanden werden, n konstante Zahlen $(\alpha_1, \ldots, \alpha_n)$ von der Eigenschaft, daß die Funktion

$$\psi(t) = \varphi(x_1^0 + \alpha_1 t, \ldots, x_n^0 + \alpha_n t)$$

ebenfalls nicht konstant ist. Setzt man nun

$$\psi_k(t) = \varphi_k(x_1^0 + \alpha_1 t, \ldots, x_n^0 + \alpha_n t),$$

so konvergieren die Funktionen $\psi_k(t)$ stetig gegen $\psi(t)$ in einer gewissen Umgebung von $t = 0$; außerdem ist nach Voraussetzung $\psi_k(t) \neq 0$ für jedes k, und nach dem soeben bewiesenen Resultat muß

$$\psi(0) = \varphi(x_1^0, \ldots, x_n^0) \neq 0$$

sein, was ja gerade zu beweisen war.

Kapitel II.

Der Kern einer Gebietsfolge.

10. Ein Abbildungssatz. In diesem Kapitel werden wir uns hauptsächlich mit Folgen von eineindeutigen topologischen Abbildungen beschäftigen. Der Satz, auf dem die ganze Theorie beruht, kann folgendermaßen ausgesprochen werden:

Satz 6. *Es sei im $2n$-dimensionalen Raum der $(x_1, \ldots, x_n)$ eine stetig konvergente Folge von analytischen Abbildungen $S_1, S_2, \ldots$ gegeben, die ein gegebenes Gebiet G topologisch transformieren. Dann ist die Grenztransformation S entweder ausgeartet, d. h. ihre Funktionaldeterminante ist identisch Null, oder S liefert ebenfalls eine topologische Abbildung von G.*

Für $n = 1$ ist bekanntlich dieser Satz eine direkte Folge des Satzes 5 (9.). Es sei nämlich

$$x' = f_k(x)$$

die Gleichung, die die Transformation S_k darstellt, und

$$x' = f(x)$$

sei die Transformationsformel für die Grenzabbildung S. Es ist zu zeigen, daß, wenn x_0 und x zwei verschiedene Punkte von G bedeuten, die Relation

$$f(x) - f(x_0) \neq 0$$

immer erfüllt ist, außer wenn $f(x)$ eine Konstante ist. Hierzu setze man

$$\varphi_k(x) = f_k(x) - f_k(x_0), \qquad \varphi(x) = f(x) - f(x_0)$$

und bemerke einfach, daß die Voraussetzungen des Satzes 5 für das punktierte Gebiet gelten, das entsteht, wenn man aus G den Punkt x_0 entfernt.

11. Der Beweis für den allgemeinen Fall ist ganz analog und nur wenig komplizierter. Nach dem Satze 1 (5.) sind die Funktionaldeterminanten Δ_k der topologischen Transformationen S_k im ganzen Gebiete G von Null verschieden, und sie konvergieren stetig in G gegen die Funktionaldeterminante Δ der Grenztransformation S. Nach dem Satze 5 (9.) muß also Δ entweder identisch gleich Null oder überall von Null verschieden sein. Ist letzteres der Fall, so muß nur noch gezeigt werden, daß die Grenztransformation S je zwei verschiedenen Punkten $(x_1, \ldots, x_n)$ und $y_1, \ldots, y_n)$ immer zwei verschiedene Punkte $(x_1', \ldots, x_n')$ bzw. $(y_1', \ldots, y_n')$ zuordnet. Durch eine auf der Hand liegende Änderung der Bezeichnungsweise kann man aber immer erreichen, daß sowohl $y_1 = \ldots = y_n = 0$ als auch $y_1' = y_2' = \ldots = y_n' = 0$ sind.

Mit dieser Normierung muß also gezeigt werden, daß die Abbildung S jeden Punkt von G, für welchen

$$|x_1|^2 + \ldots + |x_n|^2 > 0 \tag{11.1}$$

ist, in einen Punkt transformiert, für den

$$|x_1'|^2 + \ldots + |x_n'|^2 > 0 \tag{11.2}$$

ist. Da die Funktionaldeterminante Δ im Punkte $x_1 = \ldots = x_n = 0$ nicht verschwindet, ist die Abbildung S in der Umgebung dieses Punktes topologisch; es gibt dann eine Zahl $\varrho > 0$, so daß (11.2) eine Folge von (11.1) ist, solange

$$|x_1|^2 + \ldots + |x_n|^2 \leqq \varrho^2 \tag{11.3}$$

ist. Nach dem Satze 2 (7.) gibt es dann aber auch eine natürliche Zahl k_0 und eine positive Zahl σ, so daß alle Transformationen S_k, für $k \geqq k_0$, die Hyperkugel (11.3) in eine Punktmenge transformieren, welche die Hyperkugel

$$|x_1'|^2 + \ldots + |x_n'|^2 \leqq \sigma^2$$

enthält. Da die S_k das Gebiet G nach Voraussetzung topologisch abbilden, folgt für $k \geqq k_0$, daß jedem Punkte von G, der in der Punktmenge

$$|x_1|^2 + \ldots + |x_n|^2 > \varrho$$

liegt, ein Punkt entspricht, für welchen

$$|x_1'|^2 + \ldots + |x_n'|^2 \geqq \sigma^2$$

ist. Diese Eigenschaft bleibt aber beim Grenzübergang erhalten, und die behauptete Tatsache ist hiermit vollständig bewiesen.

12. Man kann auch einen Satz beweisen, der in gewisser Hinsicht eine Umkehrung des vorigen ist.

Satz 7. *Es sei die Folge der analytischen Abbildungen $S_1, S_2, \ldots$ stetig konvergent in G, und die Grenzabbildung S sei eine topologische Abbildung. Dann gibt es zu jeder abgeschlossenen Teilmenge $\overline{H}$ von G eine zugeordnete natürliche Zahl k_0, so daß alle S_k, für $k \geqq k_0$, die Punktmenge $\overline{H}$ ebenfalls topologisch abbilden.*

Wäre nämlich dieser Satz falsch, so würde es eine Folge von wachsenden Zahlen $k_1 < k_2 < \ldots$ geben und gleichzeitig zwei konvergente Folgen von Punkten $P_1, P_2, \ldots$ und $Q_1, Q_2, \ldots$ auf $\overline{H}$, so daß immer $P_m \neq Q_m$ und

$$S_{k_m} P_m = S_{k_m} Q_m \tag{12.1}$$

für $m = 1, 2, \ldots$ ist.

Es sei P_0 der Grenzpunkt von $\{P_m\}$ und Q_0 der Grenzpunkt von $\{Q_m\}$; die beiden Punkte Q_0 und P_0 liegen auf $\overline{H}$, d. h. in G. Wegen des

Satzes 2 (7.) können sie nicht zusammenfallen, weil die Grenztransformation S notwendig überall eine von Null verschiedene Funktionaldeterminante besitzt. Sie können aber auch nicht voneinander verschieden sein, weil aus (12.1) folgt

$$SP_0 = SQ_0 \tag{12.2}$$

und die Transformation S wäre, entgegen der Voraussetzung, nicht topologisch.

13. Der Unitätssatz. Es ist natürlich von äußerster Wichtigkeit, Bestimmungsstücke zu kennen, durch welche eine analytische Abbildung *eindeutig* festgelegt wird. Dieses Verlangen wird durch folgenden Satz befriedigt.

Satz 8. *Es gibt höchstens eine analytische eineindeutige Abbildung eines Gebietes G auf ein Gebiet G', bei welcher ein innerer Punkt O von G auf einen inneren Punkt O' von G' übergeht und die ersten partiellen Ableitungen der Abbildungsfunktionen im Punkte O vorgeschriebene Werte annehmen.*

Dieses wichtige Theorem ist im wesentlichen identisch mit einem Satz, den Herr Henri Cartan entdeckt und ebenso kurz wie kunstvoll bewiesen hat[8]).

Wir wählen die Bezeichnungen so, daß O bzw. O' in den Anfangspunkten der Koordinaten fallen. Dann wird eine erste Abbildung S, die G in G' transformiert, durch die Gleichungen

$$x_i' = a_{i1}x_1 + \ldots + a_{in}x_n + \ldots \qquad (i = 1, \ldots, n), \tag{13.1}$$

in denen wir nur die Glieder erster Ordnung ausgeschrieben haben, dargestellt. Jede andere Abbildung S^* von G auf G', bei der die ersten partiellen Ableitungen der Abbildungsfunktionen in O dieselben Werte wie für S haben, wird durch Gleichungen der Gestalt

$$y_i' = a_{i1}x_1 + \ldots + a_{in}x_n + \ldots \qquad (i = 1, \ldots, n) \tag{13.2}$$

vermittelt. Da die Transformation S eineindeutig sein soll, kann man aus (13.1) die x_j als Funktionen der x_i' berechnen und in (13.2) einsetzen. Man erhält, falls $y_i' \not\equiv x_i'$, Gleichungen von der Form

$$y_i' = x_i' + P_i(x_1', \ldots, x_n') + \ldots; \tag{13.3}$$

hier haben wir die homogenen Polynome niedrigster Dimension hingeschrieben, die nicht identisch Null sind.

Die Gleichungen (13.3) stellen eine Abbildung

$$T = S^* S^{-1}$$

dar, durch welche G' in sich selbst transformiert wird. Nun bemerke man, daß die Transformation T^2, durch welche G' ebenfalls auf sich selbst ab-

[8]) Henri Cartan, Les fonctions de deux variables complexes et le problème de la représentation analytique, Journ. de Math. **10** (1931), S. 1–114, insbes. théor. VII, S. 30.

gebildet wird, durch die Gleichungen

$$z_i' = x_i' + 2P_i(x_1', \ldots, x_n') + \cdots$$

dargestellt wird und daß infolgedessen T^{2^m} durch

$$w_i' = x_i' + 2^m P_i(x_1', \ldots, x_n')$$

dargestellt wird.

Nach der Bemerkung am Ende von **3.** müssen aber die verschiedenen partiellen Ableitungen von w_i' unabhängig von m gleichmäßig beschränkt sein, und man schließt hieraus, daß entgegen der Annahme

$$P_i(x_1', \ldots, x_n') \equiv 0,$$

und daß daher

$$y_i' \equiv x_i'$$

ist.

Dies ist der Beweis von Cartan, der außerdem noch bemerkt hat, daß man von der Abbildung T nur vorauszusetzen braucht, daß sie das Gebiet G' auf eine seiner Teilmengen abbildet, nicht aber, daß sie eineindeutig ist. Diese letzte Tatsache folgt übrigens direkt aus einem allgemeinen Satze, auf den wir zurückkommen werden (**22.**), und braucht uns daher hier nicht aufzuhalten.

14. Definition des Kerns. Wir wollen jetzt annehmen, daß alle Transformationen $S_1, S_2, \ldots$ topologisch sind, und daß außerdem dasselbe auch von der Grenztransformation S gilt. Unter diesen Voraussetzungen soll das transformierte Gebiet

$$G' = S\,G \tag{14.1}$$

mit Hilfe der Gebiete

$$G_k' = S_k\,G \tag{14.2}$$

charakterisiert werden.

Ist P' ein (innerer) Punkt von G', so gibt es nach dem Satze 2 (**7.**) eine Umgebung von P', die für alle hinreichend großen k in allen G_k' enthalten ist. Ist also Γ' irgendeine abgeschlossene Punktmenge, die in G' enthalten ist, so gibt es, infolge des Borelschen Überdeckungssatzes, eine natürliche Zahl k_0 und eine Umgebung $U_{\Gamma'}$ von Γ' von der Eigenschaft, daß für $k \geqq k_0$ alle G_k' die Punktmenge $U_{\Gamma'}$ als Teilmenge enthalten.

15. Wir bezeichnen jetzt mit O einen inneren Punkt von G und setzen

$$O' = S\,O. \tag{15.1}$$

Es sei ferner H' ein Gebiet, das O' in seinem Innern enthält und für hinreichend große k in allen G_k' enthalten ist. Wir wollen zeigen, daß H' ein Teilgebiet von G' ist.

Für hinreichend große k wird durch die Transformationen

$$T_k = S_k^{-1}, \tag{15.2}$$

die für jedes k in G_k' definiert sind, das Gebiet H' auf ein Teilgebiet von G topologisch transformiert. Aus der Folge der Transformationen $T_1, T_2, \ldots$ kann man also eine Teilfolge $T_{k_1}, T_{k_2}, \ldots$ aussondern, so daß alle T_{k_m} in H' definiert sind und in jedem Punkte von H' stetig gegen eine Grenztransformation T konvergieren.

Es ist zunächst zu zeigen, daß diese Grenztransformatien T nicht ausgeartet ist, und hierzu genügt es nach dem Satze 6 (**10.**) zu verifizieren, daß die Funktionaldeterminante Δ' der Transformation T im Punkte O' von Null verschieden ist.

Setzen wir

$$O_k' = S_k O, \tag{15.3}$$

so müssen nach Voraussetzung die Punkte O_k' gegen O' konvergieren. Nun hat man, wegen (15.2) und (15.3)

$$O = T_{k_m} O_{k_m}' \tag{15.4}$$

und in der Grenze, wegen der stetigen Konvergenz,

$$O = TO'. \tag{15.5}$$

Setzt man also

$$O_{k_m} = T_{k_m} O', \tag{15.6}$$

so müssen die Punkte O_{k_m} gegen O konvergieren. Nach **3.** muß also, mit Berücksichtigung der Tatsache, daß O im Inneren von G liegt, der absolute Betrag der Funktionaldeterminante Δ_{k_m} der Transformation S_{k_m} im Punkte O_{k_m} kleiner als eine feste Zahl N sein.

Nun ist nach Konstruktion, wenn man mit Δ_{k_m}' die Funktionaldeterminante von T_{k_m} im Punkte O' bezeichnet,

$$\Delta_{k_m}' \cdot \Delta_{k_m}(O_{k_m}) = 1. \tag{15.7}$$

Es ist also

$$|\Delta_{k_m}'| \geqq \frac{1}{N}, \tag{15.8}$$

und die gleiche Relation muß in der Grenze für Δ' gelten.

Durch die Grenztransformation T wird daher das Gebiet H' eineindeutig auf ein Teilgebiet H von G abgebildet.

Ist also P' irgendein Punkt von H', so ist

$$P = TP' \tag{15.9}$$

ebenfalls ein *innerer* Punkt von G. Setzt man andererseits

$$P_m = T_{k_m} P', \tag{15.10}$$

so muß P_m gegen P konvergieren, und es ist daher, wegen der stetigen Konvergenz,

$$\lim_{m=\infty} S_{k_m} P_m = S P. \tag{15.11}$$

Aus (15.10) folgt nun

$$P' = S_{k_m} P_m,$$

und die Vergleichung dieser letzten Relation mit (15.11) zeigt, daß

$$P' = S P$$

ist und daß daher jeder Punkt P' von H' und folglich daß H' selbst in G' enthalten ist.

16. Ist umgekehrt P' irgendein Punkt von G', so gibt es nach Voraussetzung einen Punkt von P von G, für welchen die Beziehung

$$P' = S P$$

gilt. Man kann dann immer ein Gebiet H finden, das sowohl P wie auch O enthält, und dessen abgeschlossene Hülle $\bar{H}$ in G enthalten ist. Das Gebiet

$$H' = S H$$

enthält dann P' und O' im Innern und wird nach **14.** für hinreichend große k durch sämtliche Gebiete G'_k überdeckt.

Um also G' zu konstruieren, kann man folgendermaßen vorgehen; man bestimme die offenen Punktmengen A'_k, die aus allen *innern* Punkten des Durchschnitts

$$G'_k \cdot G'_{k+1} \ldots$$

der Gebiete $G'_k, G'_{k+1}, \ldots$ bestehen. Hierauf setze man

$$A' = A'_1 + A'_k + \ldots;$$

G' ist dann die größte zusammenhängende, offene Punktmenge, die in A' enthalten ist und den Punkt O' enthält.

Bekanntlich kann A' auch andere Gebiete enthalten, die aber für das Problem der analytischen Abbildung nicht in Betracht kommen.

Das auf diese Weise konstruierte Gebiet G', das eindeutig definiert ist, sobald die Gebietsfolge $G'_1, G'_2, \ldots$ und der Punkt O' gegeben sind, soll der *Kern der Gebietsfolge G'_k mit dem Grundpunkt O'* genannt werden. Wir betrachten natürlich den Kern als leer nicht nur, wenn A' leer ist, sondern auch, wenn O' nicht innerer Punkt von A' ist.

Wenn eine Folge $G_1, G_2, \ldots$ und ein Punkt O gegeben sind, so sagen wir, daß die G_k gegen ihren Kern *konvergieren*, wenn alle Teilfolgen von $G_1, G_2, \ldots$ *denselben* Kern mit dem Grundpunkt O besitzen.

17. Der Hauptsatz. Wir sind jetzt in der Lage, den Hauptsatz der ganzen Theorie zu formulieren:

Satz 9. *Es sei eine Folge von topologischen Transformationen S_k gegeben, durch welche die Gebiete G_k in gewisse Gebiete G_k' übergeführt werden. Die Folge der Gebiete G_k konvergiere gegen einen nicht leeren Kern G mit dem Grundpunkt O. Im Punkte O sollen außerdem sowohl die Transformationsfunktionen wie auch ihre ersten partiellen Ableitungen konvergieren*

Sind alle diese Voraussetzungen erfüllt, so konvergiert die Folge der Abbildungen S_k dann und nur dann stetig in G gegen eine nicht ausgeartete Transformation S, wenn die Folge der Gebiete G_k' gegen einen Kern G' mit dem Grundpunkt $O' = SO$ konvergiert, der nicht leer ist.

In diesem Falle führt die Grenztransformation S den Kern G in G' über.

Zuerst bemerken wir, daß wir aus der Folge der Transformationen S_k Teilfolgen aussondern können, die im Kern G der Folge $G_1, G_2, \ldots$ stetig konvergiert. Dies sieht man am schnellsten ein, wenn man bedenkt, daß jeder Punkt von G eine Umgebung besitzt, die für hinreichend große k von allen G_k überdeckt wird.

Es sei also die Folge $S_{k_1}, S_{k_2}, \ldots$ stetig konvergent in G, und es sei außerdem die Grenztransformation S nicht ausgeartet. Dann ist die Punktmenge

$$G' = SG$$

ein Gebiet, das den Punkt $O' = SO$ in seinem Innern enthält. Jedes Gebiet H, das O in seinem Innern enthält und das mit seiner Begrenzung in G enthalten ist, wird durch S in eine Punktmenge H' transformiert, die im Kern der Gebietsfolge $G_{k_1}', G_{k_2}', \ldots$ mit dem Grundpunkt O' liegt. Also muß G' auch in diesem Kern enthalten sein.

18. Ist umgekehrt H' ein Gebiet, das O' in seinem Inneren enthält und mit seiner Begrenzung im soeben erwähnten Kern enthalten ist, so wird für hinreichend große m das Gebiet H' von sämtlichen G_{k_m}' überdeckt. Also wird H' für dieselben Werte von m durch die Transformationen $S_{k_m}^{-1}$ in Teilgebiete von G_{k_m} übergeführt, und man schließt hieraus, wie in **15.**, daß H' durch S^{-1} in eine Teilmenge von G transformiert wird.

Das Gebiet G' muß demnach mit dem Kern der Gebietsfolge $G_{k_1}', G_{k_2}', \ldots$, der den Grundpunkt O' hat, identisch sein.

Konvergiert nun die ursprüngliche Folge der Transformationen $S_1, S_2, \ldots$ gegen eine nicht ausgeartete Transformation S, so konvergieren alle Teilfolgen $S_{k_1}, S_{k_2}, \ldots$ gegen dieselbe Grenztransformation S, und die Gebietsfolgen $G_{k_1}', G_{k_2}', \ldots$ müssen daher immer denselben Kern G' mit dem Grundpunkt O' besitzen. Die Gebietsfolge $G_1', G_2', \ldots$ muß also gegen diesen Kern konvergieren.

Ist umgekehrt letzteres der Fall und der Kern G' mit dem Grundpunkt O' nicht leer, so kann man, genau wie in **15.**, eine endliche obere Schranke für den absoluten Betrag der Funktionaldeterminanten Δ_k' aller inversen Transformationen S_k^{-1} im Punkte O' berechnen, die in O' definiert sind. Hieraus schließt man dann, daß jede konvergente Teilfolge $S_{k_1}, S_{k_2}, \ldots$ der Transformationen S_k gegen eine nicht ausgeartete Transformation S konvergiert, durch welche der Kern G auf den Kern G' abgebildet wird. Wegen des Unitätssatzes von **13.** ist aber diese Grenztransformation S unabhängig von der Wahl der Teilfolge S_{k_m}, und hieraus folgt in bekannter Weise die Konvergenz der ursprünglichen Folge $S_1, S_2, \ldots$.

19. Der soeben bewiesene Satz ist besonders dann gut zu gebrauchen, wenn man von vornherein feststellen kann, daß die Gebietsfolge $G_1', G_2', \ldots$ gegen ihren Kern konvergiert.

Es ist übrigens sehr leicht aus einer beliebigen Folge von Gebieten

$$G_1^*, G_2^*, \ldots \tag{19.1}$$

eine Teilfolge auszusondern, die gegen ihren Kern konvergiert. Hierzu betrachten wir eine Folge von Punkten

$$Q_1, Q_2, \ldots, \tag{19.2}$$

die den $2n$-dimensionalen Raum überall dicht besetzt. Es sei Q_{m_1} der erste Punkt von (19.2), der mit einer seiner Umgebungen in unendlich vielen Gebieten (19.1) enthalten ist. Man bezeichne mit ϱ_1 den Limes superior der Radien aller Hyperkugeln mit Q_{m_1} als Mittelpunkt, die in einem der Gebiete G_k^* enthalten ist. Ferner sei $\varrho_{11}, \varrho_{12}, \ldots$ eine monoton wachsende Folge von positiven Zahlen, die gegen ϱ_1 konvergiert. Wir bezeichnen mit $\varkappa_{1j}$ bzw. mit $\varkappa_1$ die Hyperkugel mit dem Mittelpunkte Q_{m_1} und dem Radius ϱ_{1j} bzw. ϱ_1. Dann können wir aus (19.1) eine Teilfolge

$$G_{11}, G_{12}, \ldots \tag{19.3}$$

auswählen, von der Eigenschaft, daß G_{1j} immer die Hyperkugel $\varkappa_{1j}$ enthält. Wir iterieren dieses Verfahren, indem wir die Folge (19.1) durch die Folge

$$G_{12}, G_{13}, \ldots, \tag{19.4}$$

und die Folge (19.2) durch

$$Q_{m_1+1}, Q_{m_2+2}, \ldots \tag{19.5}$$

ersetzen und auf ähnliche Weise unbegrenzt fortfahren. Wir erhalten unendlich viele Gebietsfolgen

$$G_{pp}, G_{p(p+1)}, \ldots \tag{19.6}$$

von denen jede eine Teilfolge der vorhergehenden ist, und auf welche wir das Diagonalverfahren anwenden können.

Hat die ursprüngliche Folge (19.1) einen nicht verschwindenden Kern mit dem Grundpunkt O, so gilt dasselbe von der Teilfolge

$$G_{11}, G_{22}, \ldots, \tag{19.7}$$

die außerdem gegen ihren Kern konvergieren muß.

Kapitel III.

Die analytischen Selbstabbildungen eines Gebietes.

20. Ein Konvergenzsatz. Einige Resultate, die sonst recht umständlich zu beweisen wären, werden fast selbstverständlich, wenn man den Satz 9 (17.) benutzt. Betrachten wir z. B. eine konvergente Folge von analytischen Abbildungen $S_1, S_2, \ldots$, durch welche ein Gebiet G auf sich selbst topologisch abgebildet wird.

Wir nehmen an, daß ein beliebiger innerer Punkt P von G durch die Transformation S_k in Punkten

$$P_k' = S_k P \tag{20.1}$$

abgebildet wird, die gegen einen Punkt P' konvergieren, der ebenfalls im *Inneren* von G liegt. Insbesondere ist dies immer der Fall, wenn die Grenztransformation S nicht ausgeartet ist; denn die offene Punktmenge

$$G' = SG \tag{20.2}$$

ist jedenfalls in der abgeschlossenen Hülle $\bar{G}$ von G enthalten, und P' ist als innerer Punkt von G auch innerer Punkt von $\bar{G}$.

Nun hat man nach Voraussetzung mit der Bezeichnung des vorigen Kapitels

$$G_k' = S_k G = G. \tag{20.3}$$

Die Folge der Gebiete G_k' konvergiert gegen einen nicht leeren Kern mit dem Grundpunkt P', der mit dem Gebiete G zusammenfällt.

Nach dem Satz 9 (17.) ist also erstens die Tatsache, daß P' im Inneren von G liegt, hinreichend dafür, daß die Grenztransformation S nicht ausgeartet sei, und zweitens ist dann immer

$$G' = SG = G,$$

d. h. die Grenztransformation S liefert in diesem Fall eine Selbstabbildung von G.

Es gilt also, wenn wir die vorhergehenden Schlüsse zusammenfassen, der

Satz 10. *Jede konvergente Folge von eineindeutigen analytischen Abbildungen eines beschränkten Gebietes G auf sich selbst konvergiert entweder gegen eine ebensolche Abbildung oder aber gegen eine ausgeartete Transformation, bei welcher jeder innere Punkt von G in einen Randpunkt dieses Gebietes übergeführt wird.*

21. Die Menge der iterierten Selbsttransformationen. Wir betrachten wieder eine analytische Transformation S, durch welche ein beschränktes Gebiet G eineindeutig in sich selbst transformiert wird. Dann haben die iterierten Transformationen $S^2, S^3, \ldots$ dieselbe Eigenschaft.

Es sei P irgendein innerer Punkt von G; wir bezeichnen mit $P_1, P_2, \ldots$ die transformierten P durch die Abbildungen $S, S^2, \ldots$. Es kann vorkommen, daß *alle* Häufungspunkte der Punktmenge $\{P_k\}$ auf der Begrenzung von G liegen. Dann werden nach dem Satz des vorigen Paragraphen alle konvergente Teilfolgen $S^{m_1}, S^{m_2}, \ldots$ der Folge $S, S^2, \ldots$ gegen ausgeartete Transformationen konvergieren.

Gibt es dagegen einen Häufungspunkt P_0 der Punktmenge $\{P_k\}$, der im Inneren von G liegt, so kann man mindestens eine konvergente Folge von Transformationen

$$S^{m_1}, S^{m_2}, \ldots \tag{21.1}$$

finden, deren Grenztransformation T das Gebiet G in sich selbst überführt. Man kann die Folge (21.1) immer so wählen, daß

$$m_1 < m_2 < \ldots \tag{21.2}$$

ist. Wir setzen

$$p_j = m_{j+1} - m_j; \tag{21.3}$$

es ist immer möglich, aus der Folge der Transformationen

$$S^{p_1}, S^{p_2}, \ldots$$

eine Teilfolge

$$S^{p_{k_1}}, S^{p_{k_2}}, \ldots \tag{21.4}$$

auszuwählen, die gegen eine Grenztransformation E stetig konvergiert. Wir wollen beweisen, daß E die identische Transformation ist.

Es sei nämlich Q ein beliebiger innerer Punkt von G; da die Grenztransformation T der Folge (21.1) nicht ausgeartet ist, gibt es im Inneren von G einen Punkt P, der die Gleichung

$$Q = TP \tag{21.5}$$

befriedigt. Setzt man also

$$P_{m_j} = S^{m_j} P, \tag{21.6}$$

so hat man

$$P_{m_j} \to Q. \tag{21.7}$$

Nun ist aber wegen (21.3)

$$P_{m_{k_j+1}} = S^{p_{k_j}} P_{m_{k_j}}; \tag{21.8}$$

aus

$$S^{p_{k_j}} \to E \tag{21.9}$$

folgt dann, mit Berücksichtigung von (21.7)

$$Q = EQ, \tag{21.10}$$

womit bewiesen ist, daß E die identische Transformation darstellt.

Es gilt also der

Satz 11. *Es sei S eine Transformation, die ein Gebiet G eineindeutig auf sich selbst abbildet. Man kann immer aus der Folge*

$$S^2, S^3, \ldots \tag{21.11}$$

der iterierten Transformationen eine konvergente Teilfolge auswählen, die gegen die identische Transformation konvergiert, außer für den Fall, daß sämtliche konvergente Teilfolgen von (21.11) *gegen ausgeartete Transformationen konvergieren.*

Bemerkung. Kommt insbesondere in der Folge (21.3) eine und dieselbe Zahl unendlich oft vor, ist z. B. immer $p_{j_k} = p_0$ für $k = 1, 2, \ldots$, so ist schon

$$S^{p_0} = E.$$

22. Ein Satz von Henri Cartan. Wir betrachten ein Gebiet G, das den Punkt O in seinem Inneren enthält, und eine analytische Abbildung S, durch welche G in eine Punktmenge G_1 transformiert wird, die in G enthalten ist und die außerdem O zum Fixpunkt besitzt. Es soll also sein

$$G_1 = SG < G, \quad O = SO. \tag{22.1}$$

Bezeichnet man mit α den absoluten Betrag der Funktionaldeterminante Δ der Transformation S im Punkte O, so ist der absolute Betrag der Funktionaldeterminante der iterierten Transformation S^k im Punkte O gleich α^k. Nach **3.** gibt es aber eine endliche Schranke N, so daß für jedes k

$$\alpha^k < N$$

ist. Hieraus folgt dann ohne weiteres

$$\alpha \leqq 1. \tag{22.1}$$

Bildet die Transformation S das Gebiet G eineindeutig und umkehrbar stetig auf sich selbst ab, so findet man, wenn man dieselbe Schlußweise auf die Transformation S^{-1} anwendet, daß auch

$$\frac{1}{\alpha} \leqq 1$$

ist. Hieraus folgt dann

$$\alpha = 1. \tag{22.2}$$

Herr Cartan hat nun gefunden, daß auch die Umkehrung dieses Satzes gilt, d. h. daß aus der Relation (22.2) folgt, daß S eine topologische Selbst-

abbildung des Gebietes G sein muß[9]). Den Beweis, den er für dieses Resultat gegeben hat und der auf den vierdimensionalen Raum zugeschnitten ist, kann man nun durch Benutzung der vorhergehenden Sätze wesentlich vereinfachen.

23. Wir wollen zuerst den Cartanschen Satz für den speziellen Fall beweisen, daß S eine *eineindeutige topologische* Abbildung ist, die das Gebiet G in ein Teilgebiet

$$G_1 = SG \tag{23.1}$$

transformiert. Es ist dann zu zeigen, daß aus $\alpha = 1$ die Relation $G_1 = G$ folgt.

Setzt man

$$G_{k+1} = SG_k \qquad (k = 1, 2, \ldots), \tag{23.2}$$

so sind die Gebiete $G_1, G_2, \ldots$ ineinandergeschachtelt; es gilt

$$G_1 \succ G_2 \succ \ldots \tag{23.3}$$

und außerdem wegen (23.1) und (23.2)

$$G_k = S^k G. \tag{23.4}$$

Im Punkte O ist der absolute Betrag der Funktionaldeterminante Δ_k der Transformation S^k immer gleich $\alpha^k = 1$. Wählt man also aus der Folge der Transformation $S, S^2, \ldots$ eine konvergente Teilfolge

$$S^{k_1}, S^{k_2}, \ldots \tag{23.5}$$

aus, die gegen eine Grenztransformation T konvergiert, so ist die Abbildung T nach dem Satze 6 (**10.**) nicht ausgeartet. Durch T wird das Gebiet G wieder in ein Gebiet

$$G_0 = TG \tag{23.6}$$

transformiert; hierbei ist G_0 der Kern der Gebietsfolge

$$G_{k_1}, G_{k_2}, \ldots \tag{23.7}$$

mit dem Grundpunkt O (Satz 9 (**17.**)), und wegen (23.3) ist

$$G_0 = G_1 G_2 \ldots \tag{23.8}$$

der gemeinsame Durchschnitt aller Gebiete G_k. Nun ist mit der Folge (23.5) auch die Folge

$$S^{k_1+1}, S^{k_2+1}, \ldots \tag{23.9}$$

konvergent, und die Grenztransformation lautet jetzt TS. Nach den letzten Schlüssen muß also auch

$$TSG = G_0 \tag{23.10}$$

[9]) Zitat unter [7]), insbesondere Theorem II, S. 201.

sein. Aus (23.6) und (23.10) folgt nunmehr

$$TG = TSG,$$

und man entnimmt hieraus

$$T^{-1}TG = T^{-1}TSG,$$

und daher auch

$$G = SG,$$

was ja gerade zu beweisen war.

24. Jetzt können wir den Satz von Cartan allgemein beweisen. Aus (22.2) folgt, daß eine gewisse Umgebung U_0 des Punktes O durch die Transformation S topologisch abgebildet wird.

Wir bezeichnen, ähnlich wie in 8., mit $G(\varrho)$ die Gesamtheit der Punkte von G, die man mit O durch eine Kurve γ, deren metrische Länge $L_G(\gamma) < \varrho$ ist, verbinden kann. Wählt man die positive Zahl ϱ hinreichend klein, so ist

$$G(\varrho) \prec U_0. \tag{24.1}$$

Nach dem Satze 4 (8.) wird durch die Transformation S das Gebiet $G(\varrho)$ in eine Punktmenge

$$G_1(\varrho) = SG(\varrho) \tag{24.2}$$

transformiert, die im Inneren von $G(\varrho)$ liegt. Wegen (24.1) ist diese Abbildung topologisch, und nach dem Resultate des letzten Paragraphen ist also

$$G(\varrho) = SG(\varrho). \tag{24.3}$$

Hieraus folgt, daß die analytische Transformation S eine topologische Selbstabbildung des Gebietes $G(\varrho)$ liefert, und wir können den Satz 11 (**21.**) anwenden. Es gibt danach, weil O ein Fixpunkt ist, eine Teilfolge

$$S^{p_1}, S^{p_2}, \ldots \tag{24.3}$$

der Folge der iterierten Transformationen $S, S^2, S^3, \ldots$, die gegen die identische Transformation E konvergiert, und diese Konvergenz gilt, wegen des Vitalischen Satzes, nicht nur im Inneren des Gebietes $G(\varrho)$, sondern im ganzen Gebiete G.

Es ist nun zuerst zu zeigen, daß die Funktionaldeterminante $\Delta(P)$ der Transformation S in jedem Punkte P von G von Null verschieden ist[10]). Für hinreichend große j ist sicher die Funktionaldeterminante $\Delta_{p_j}(P)$ im Punkte P von Null verschieden. Nun gilt aber, falls man

$$P_k = S^k P \tag{24.4}$$

[10]) Dieser Teil des Beweises fehlt bei Herrn Cartan, der sich offenbar auf die Resultate gestützt hat, die in der Fußnote [4]) (5) erwähnt werden.

setzt, die Formel

$$\Delta_{p_j} P = \Delta(P) \cdot \prod_{k=1}^{p_j - 1} \Delta(P_k), \tag{24.5}$$

woraus die gewünschte Relation

$$\Delta(P) \neq 0 \tag{24.6}$$

unmittelbar zu entnehmen ist.

Um nun zu zeigen, daß die Abbildung von G durch S schlicht ist, muß man nur noch feststellen, daß, wenn P und Q zwei verschiedene Punkte von G sind, niemals

$$SP = SQ \tag{24.7}$$

sein kann. Aus dieser letzten Relation würde aber ganz allgemein für jedes k folgen

$$S^k P = S^k Q,$$

und dieses ist unverträglich mit den Relationen

$$\lim_{j=\infty} S^{p_j} P = P, \qquad \lim_{j=\infty} S^{p_j} Q = Q,$$

die aus unseren Annahmen folgen.

Durch die Transformation S wird also das Gebiet G topologisch und schlicht auf eins seiner Teilgebiete abgebildet, und nach **23.** muß daher S eine Selbstabbildung von G sein.

Der Satz, den wir hiermit bewiesen haben, kann folgendermaßen ausgesprochen werden:

Satz 12. *Es sei S eine Transformation mit dem Fixpunkt O, durch welche ein Gebiet G in eine Punktmenge G' transformiert wird, die im Inneren von G liegt. Dann ist der absolute Betrag der Funktionaldeterminante Δ der Transformation S nie größer als Eins und dann und nur dann gleich Eins, wenn S eine eineindeutige Selbstabbildung von G darstellt.*

25. Herr Cartan hat in der zuletzt erwähnten Arbeit einige Folgerungen des letzten Satzes abgeleitet, die ich noch besprechen möchte, nicht nur weil ihr Beweis mit Hilfe der vorhergehenden Theorie sehr einfach ist, sondern vor allem, weil man sie in einem Punkte vervollständigen kann.

Wir wollen folgenden Satz beweisen:

Satz 13. *Es sei H ein Gebiet, das den Punkt O in seinem Inneren enthält und dessen abgeschlossene Hülle $\bar{H}$ im beschränkten Gebiet G enthalten ist. Wir betrachten die analytischen Transformationen S, die den Bedingungen*

$$G' = SG \prec G, \quad O' = SO = O \tag{25.1}$$

genügen, und bezeichnen mit Δ die Funktionaldeterminante von S im Punkte O.

Dann gibt es erstens eine positive Zahl $\omega' < 1$, so daß für alle Transformationen S, für welche $|\Delta| > \omega'$ ist, die Abbildung

$$H' = SH \tag{25.2}$$

topologisch ist. Und zweitens gibt es eine Zahl $\omega'' < 1$, so daß, sobald $|\Delta| > \omega''$ ist, das Gebiet H das topologische Bild eines Teilgebietes von G sein muß.

Wir betrachten eine monotone Folge von positiven Zahlen

$$\alpha_1 < \alpha_2 < \ldots, \tag{25.3}$$

die der Bedingung

$$\lim_{k=\infty} \alpha_k = 1 \tag{25.4}$$

genügen. Entweder gibt es eine Zahl α_{k_0}, so daß der erste Teil unseres Satzes richtig ist, falls man

$$\omega' = \alpha_{k_0}$$

setzt, oder man kann zu jedem α_k eine Transformation S_k mit der Bedingung

$$|\Delta_k| > \alpha_k$$

zuordnen, für welche die Transformation

$$H_k' = S_k H$$

nicht topologisch ist. Aus der so konstruierten Folge $S_1, S_2, \ldots$ wählen wir eine konvergente Teilfolge $S_{k_1}, S_{k_2}, \ldots$ aus. Für diese letzten Transformationen hat man

$$\lim_{m=\infty} |\Delta_{k_m}| = 1.$$

Nach dem Satze 12 muß also die Grenztransformation S eine topologische Selbstabbildung von G sein. Nach dem Satze 7 (12.) muß also, entgegen der Voraussetzung, für hinreichend große Werte von m das Gebiet H durch S_{k_m} topologisch abgebildet werden.

Ähnlich kann man schließen, daß, wenn der zweite Teil des Satzes 13 nicht bestehen sollte, eine konvergente Folge von Transformationen $S_1, S_2, \ldots$ konstruiert werden kann, deren Funktionaldeterminanten Δ_k im Punkte O der Bedingung

$$\lim_{k=\infty} |\Delta_k| = 1$$

genügen, und von der Eigenschaft, daß kein einziges S_k Teilgebiete von G topologisch in H transformiert. Nach dem Satze 12 (24.) muß wiederum die Grenztransformation S eine topologische Selbstabbildung von G sein. Man kann daher setzen

$$\overline{K} = S^{-1}\overline{H};$$

ferner kann man eine Umgebung U der abgeschlossenen Punktmenge $\overline{K}$ konstruieren, deren abgeschlossene Hülle $\overline{U}$ in G enthalten ist. Nach dem Satz 7 (**12.**) ist dann für hinreichend große k die Abbildung

$$U_k = S_k U$$

topologisch. Durch die Grenztransformation S wird U auf ein Gebiet SU abgebildet, das die abgeschlossene Punktmenge $\overline{H}$ enthält. Nach **14.** gibt es also eine Zahl k_0, so daß für $k > k_0$ die Punktmenge $S_k U$ ebenfalls das abgeschlossene Gebiet $\overline{H}$ in ihrem Inneren enthalten muß. Dies widerspricht aber den Eigenschaften, die für die Transformationen S_k festgesetzt worden waren. Hiermit ist der angekündigte Satz vollständig bewiesen.

26. Über die konforme Abbildung mehrfach zusammenhängender ebener Gebiete. Das Interesse, das die Sätze 12 und 13 beanspruchen, wird durch die Tatsache beeinträchtigt, daß es viele Gebiete gibt, die eine viel einschneidendere Eigenschaft besitzen, so daß für diese Gebiete die Cartanschen Resultate gegenstandslos werden.

Wir betrachten z. B. in der Ebene ein zweifach zusammenhängendes Gebiet, das man auf einen Kreisring, z. B. auf

$$r < |x - 1| < R \qquad (r < 1,\ R > 1) \tag{26.1}$$

abbilden kann. Das Gebiet (26.1), das wir wieder mit G bezeichnen, enthält den Nullpunkt in seinem Inneren.

Wir betrachten die normale Familie der analytischen Funktionen $f(x)$, die in G eindeutig und regulär sind und außerdem die beiden Bedingungen

$$f(0) = 0, \quad r < |f(x) - 1| < R \text{ für } x \prec G \tag{26.2}$$

befriedigen.

Ist nun $f(x)$ irgendeine nicht konstante Funktion unserer Familie und bezeichnen wir wieder mit S die Transformation

$$x' = f(x), \tag{26.3}$$

so stellt die Punktmenge

$$G' = S\,G$$

eine Riemannsche Fläche dar, die nach Voraussetzung auf einem Teil unseres Kreisringes ausgebreitet ist.

Nun kann man jeder rektifizierbaren Kurve c, die in G verläuft, eine nichteuklidische Länge zuordnen und diese Metrik auf das Gebiet G' übertragen.

Ist dann

$$c' = S\,c \tag{26.4}$$

das Bild von c und bezeichnet man mit $L_G(c)$ bzw. mit $L_{G'}(c')$ die nichteuklidischen Längen der beiden Kurven in den beiden schon erwähnten

Metriken, so gelten die Beziehungen[11])

$$L_{G'}(c') = L_G(c), \tag{26.5}$$

$$L_G(c') \leqq L_G(c), \tag{26.6}$$

von denen die zweite eine Verallgemeinerung des Schwarzschen Lemmas darstellt. In der letzten Gleichung gilt das Gleichheitszeichen dann und nur dann, wenn die Transformation S eine eineindeutige Selbstabbildung von G liefert, die den Anfangspunkt O fest läßt.

Ist in (26.1) das Produkt $Rr \neq 1$, so ist bekanntlich die identische Transformation die einzige derartige Selbstabbildung; ist $Rr = 1$, so gibt es noch eine zweite, bei welcher die beiden Ränder des Kreisringes G miteinander vertauscht werden.

Wir bezeichnen nun mit $\varkappa$ den Kreis

$$|x - 1| = 1$$

und nehmen von unserer Transformation S an, daß sie den Kreis $\varkappa$ in eine topologisch äquivalente Kurve

$$\varkappa' = S\varkappa$$

überführt. Ist dann γ irgendeine geschlossene Kurve in G, die den inneren Rand des Ringes einmal umkreist, also dem Kreise $\varkappa$ topologisch äquivalent ist, so gilt dasselbe vom Bilde

$$\gamma' = S\gamma$$

dieser Kurve.

Die nichteuklidische Länge $L_G\gamma$ einer solchen Kurve γ, wenn man sie genau einmal durchläuft, erreicht nun ihren kleinsten Wert, wenn γ mit der Symmetrielinie c_0 von G, d. h. mit dem Kreis

$$|x - 1| = \sqrt{Rr}$$

zusammenfällt. Nun ist $c_0' = Sc_0$ der Kurve c_0 topologisch äquivalent und daher

$$L_G c_0' \geqq L_G c_0. \tag{26.7}$$

Die Vergleichung von (26.6) mit der letzten Relation zeigt, daß $L_G c_0' = L_G c_0$ sein muß, so daß jede Abbildung S unserer Familie, bei welcher die Kurve $\varkappa' = S\varkappa$ dem Kreise $\varkappa$ topologisch äquivalent ist, eine topologische Selbstabbildung von G sein muß.

Zu demselben Ergebnis wird man geführt, wenn das ursprünglich gegebene zweifach zusammenhängende Gebiet nicht auf einen Kreisring, wohl aber auf eine punktierte Kreisscheibe abbildbar ist. Die Schlußweise muß

---

[11]) C. Carathéodory, Über eine Verallgemeinerung der Picardschen Sätze, Berl. Sitzungsber. 1920, S. 202—209, insbes. § 4, S. 204. Ähnliche Gedankengänge sind später wiederholt von P. Koebe und anderen benutzt worden.

aber dann ein wenig modifiziert werden. In diesem Falle kann man für G das Gebiet

$$0 < |x-1| < R \qquad (R>1) \tag{26.8}$$

wählen. Nach dem Riemannschen Satze sind dann alle Funktionen $f(x)$ unserer Familie auch im Punkte $x=1$ regulär. Jede Abbildung S, bei welcher der Kreis $\varkappa$ in eine topologisch äquivalente Kurve übergeführt wird, muß dann durch eine Funktion $f(x)$ erzeugt werden, für welche

$$f(1)=1$$

ist. Die Transformation S muß also die beiden Fixpunkte $x=0$ und $x=1$ besitzen, und das Schwarzsche Lemma lehrt uns dann, daß S mit der identischen Transformation zusammenfällt.

27. Man bezeichne nun mit Ω die obere Grenze von $|f'(0)|$ für alle Transformationen unserer Familie, bei welchen der Kreis $\varkappa$ in eine Kurve $\varkappa' = S\varkappa$ übergeht, die dem Kreise $\varkappa$ nicht topologisch äquivalent ist. Eine Schlußweise, die derjenigen des § 25 analog ist, zeigt dann, daß in jedem Falle $\Omega < 1$ ist.

Dieses Resultat kann man mit Hilfe von geringfügigen Ergänzungen der Beweisanordnung auf Gebiete von beliebigem Zusammenhang übertragen. Es gilt der

Satz 14. *Es sei G ein beschränktes mehrfach zusammenhängendes Gebiet der Ebene, das den Punkt $x=0$ in seinem Inneren enthält. Die Funktion $f(x)$ sei regulär und eindeutig in G; es sei außerdem $f(0)=0$, und es werde durch $x'=f(x)$ das Gebiet G auf eine Punktmenge abgebidet, die in G enthalten ist. Dann gibt es eine positive Zahl $\Omega<1$, so daß jede Funktion $f(x)$, die die angegebenen Eigenschaften besitzt, eine topologische Selbstabbildung von G liefert, sobald*

$$|f'(0)| > \Omega$$

ist.

Nur wenn der Punkt $x=0$ ein Symmetriepunkt von G ist, kann außerdem die erwähnte topologische Selbstabbildung von G von der identischen Transformation verschieden sein.

Ist das Gebiet G nicht beschränkt, so kann man Beispiele angeben, bei welchen der vorige Satz sogar mit $\Omega=0$ gilt. Dies findet insbesondere statt, wenn man für G die ganze Ebene wählt, aus welcher man die Punkte -1, $+1$, ∞ entfernt hat.

Der Satz 14 zeigt, daß die Starrheit von mehrfach zusammenhängenden ebenen Gebieten gegenüber konformen Abbildungen viel größer ist, als gewöhnlich angenommen wird. Es ist zu erwarten, daß ähnliche Verhältnisse in mehrdimensionalen Räumen stattfinden, nicht nur, wie selbstver-

ständlich, für Gebiete mehrfachen Zusammenhanges, sondern unter Umständen auch für solche Gebiete, die den topologischen Zusammenhang der Hyperkugel haben. In der Tat zeigen alle unsere bisherigen Erfahrungen über analytische Abbildungen mehrdimensionaler Gebiete, daß schon ganz einfache Kreiskörper eher den mehrfach- als den einfachzusammenhängenden ebenen Gebieten gleich zu setzen sind.

Kapitel IV.

Die Maximalteiler eines Gebietes.

28. Definition. Wir betrachten zwei beschränkte Gebiete G und H', von denen das erste im Raume der $(x_1, \ldots, x_n)$, das zweite aber im Raume der $(x_1', \ldots, x_n')$ enthalten sein soll. Beide Gebiete sollen den Anfangspunkt der Koordinaten O und O' der Räume, in denen sie sich befinden, in ihrem Inneren enthalten.

Wir betrachten die Gesamtheit der Transformationen S', die H' topologisch in ein Teilgebiet H von G und O' in O überführen. Es soll also sein

$$H = S'H' \prec G, \qquad O = S'O'. \tag{28.1}$$

Nach **3.** sind die Funktionaldeterminanten Δ' aller dieser Transformationen, wenn man sie für den Punkt O' berechnet, ihrem absoluten Betrage nach kleiner als eine feste Schranke.

Man kann daher eine konvergente Folge $S_1', S_2', \ldots$ von Transformationen innerhalb der betrachteten Familie finden, für welche die absoluten Beträge $|\Delta_1'|, |\Delta_2'|, \ldots$ der zugehörigen Funktionaldeterminanten in O' gegen ihre obere Grenze α konvergieren.

Durch die Grenztransformation S^* wird dann nach dem Satze 6 (**10.**) das Gebiet H' in ein Teilgebiet H^* von G übergeführt, das wir einen Maximalteiler von G nennen wollen.

29. Eigenschaften der Maximalteiler. Es sei H^* der Maximalteiler von G, den wir soeben konstruiert haben; man bezeichne mit S eine beliebige analytische Transformation mit dem Fixpunkte O, durch welche das Gebiet H^* in ein Teilgebiet H von G übergeführt wird. Ist dann Δ die Funktionaldeterminante von S im Punkte O, so hat die Funktionaldeterminante der Transformation SS^* in O' den Wert $\Delta\Delta^*$. Nach **28.** muß also

$$|\Delta| \cdot |\Delta^*| \leqq \alpha$$

sein. Da aber nach unserer Konstruktion $|\Delta^*| = \alpha$ war, finden wir

$$|\Delta| \leqq 1, \tag{29.1}$$

und man sieht leicht ein, daß, wenn die Bedingung (29.1) für alle Trans-

formationen S besteht, die betrachtet worden sind, das Gebiet H^* ein Maximalteiler von G sein muß.

Wir können demnach folgende Definition aufstellen:

Definition. *Es sei G ein beschränktes Gebiet, das den Anfangspunkt O der Koordinaten in seinem Inneren enthält, und H^* ein Teilgebiet von G, das dieselbe Eigenschaft besitzt. Das Gebiet H^* heißt ein Maximalteiler von G, wenn jede analytische Transformation S mit dem Fixpunkt O, die H^* topologisch auf ein Teilgebiet von G abbildet, im Punkte O eine Funktionaldeterminante besitzt, deren absoluter Betrag nie größer als Eins ist.*

Nach dem Satze 12 (**24.**) ist das Gebiet G selbst unter seinen Maximalteilern enthalten.

30. Aus der letzten Bemerkung folgt, daß man nie das Gebiet G auf einen seiner Maximalteiler H^* abbilden kann, wenn der Anfangspunkt O sich selbst entsprechen soll, außer wenn H^* mit G zusammenfällt. Falls nämlich eine derartige Abbildung möglich wäre, müßte, wegen des Satzes 12 (**24.**), die inverse Abbildung von H^* auf G, entgegen der Definition des Maximalteilers, eine Funktionaldeterminante im Punkte O besitzen, deren absoluter Betrag >1 wäre.

Es sei H^{**} ein Teilgebiet von G, das den Maximalteiler H^* in seinem Inneren enthält.

Es ist erstens selbstverständlich, daß H^* ein Maximalteiler von H^{**} sein muß, denn jede Transformation S mit dem Fixpunkt O, die H^* in ein Teilgebiet von H^{**} überführt, muß in O eine Funktionaldeterminante Δ besitzen, deren absoluter Betrag $\leqq 1$ ist. Man sieht aber auch leicht ein, daß H^{**} ebenfalls ein Maximalteiler von G sein muß. Denn bei jeder der betrachteten Abbildungen, die H^{**} auf ein Teilgebiet von G transformiert, wird erst recht H^* in ein solches übergeführt. Gibt man sich also eine Schar von Gebieten $H(\lambda)$, die von einem positiven Parameter λ abhängen und den Bedingungen

$$(30.1) \qquad H(0) = H^*, \quad H(\lambda') \prec H(\lambda'') \text{ für } \lambda' < \lambda'', \quad H(\lambda) \prec G$$

genügen, so ist für jedes positive λ das Gebiet $H(\lambda)$ Maximalteiler von G, und $H(\lambda')$ ist Maximalteiler von $H(\lambda'')$, sobald $\lambda' < \lambda''$ ist.

Auf diese Weise kann man kontinuierliche Scharen von Gebieten konstruierén, von denen keine zwei mit Erhaltung des Nullpunktes aufeinander abbildbar sind.

31. Die vorigen kurzen Andeutungen zeigen, daß der Begriff des Maximalteilers in vielen Fällen brauchbar ist. Von Interesse dürfte noch folgende Bemerkung sein.

Das Gebiet H', das wir in 28. betrachtet haben, möge monoton und stetig von einem Parameter λ abhängen. Es soll also erstens

$$H'(\lambda_1) < H'(\lambda_2) \text{ für } \lambda_1 < \lambda_2 \tag{31.1}$$

sein, und zweitens soll, wenn P irgendeinen inneren Punkt von $H'(\lambda)$ bedeutet, für hinreichend kleine positive ε das Gebiet $H'(\lambda - \varepsilon)$ ebenfalls den Punkt P enthalten.

Wir bezeichnen mit $\alpha(\lambda)$ die obere Grenze der absoluten Beträge der Funktionaldeterminanten Δ' im Punkte O', für alle Transformationen S', die $H'(\lambda)$ in ein Teilgebiet von G und O' in O überführen.

Selbstverständlich ist $\alpha(\lambda)$ monoton abnehmend, also

$$\alpha(\lambda_1) \geqq \alpha(\lambda_2) \text{ für } \lambda_1 < \lambda_2. \tag{31.2}$$

Betrachtet man nun eine monoton wachsende konvergente Folge von Parameterwerten λ_k, so folgt aus

$$\lambda_k < \lambda_{k+1}, \quad \lim_{k=\infty} \lambda_k = \lambda_0 \tag{31.3}$$

die Relation

$$\alpha(\lambda_0) \leqq \lim_{k=\infty} \alpha(\lambda_k).$$

Nun konvergieren die Gebiete $H'(\lambda_k)$ gegen ihren Kern $H'(\lambda_0)$; nach dem Satze 9 (17.) gibt es daher eine analytische Abbildung von $H'(\lambda_0)$ auf ein Teilgebiet von G, bei welcher O' in O übergeführt wird und der absolute Betrag der Funktionaldeterminante in O' den Wert $\lim_{k=\infty} \alpha(\lambda_k)$ besitzt. Es muß daher auch

$$\alpha(\lambda_0) \geqq \lim_{k=\infty} \alpha(\lambda_k)$$

sein, und es folgt hieraus, *daß die monotone Funktion $\alpha(\lambda)$ halbstetig nach oben ist.*

Dagegen braucht $\alpha(\lambda)$ nicht stetig zu sein, wie man an Beispielen mit zweifach zusammenhängenden ebenen Gebieten leicht verifizieren kann.

32. Hyperkugel und Polyzylinder. Wir wollen einen Maximalteiler H^* von G bestimmen für den Fall, daß G mit dem Polyzylinder

$$|x_1| < 1, \ldots, |x_n| < 1 \tag{32.1}$$

zusammenfällt und H' die Hyperkugel

$$u_1 \bar{u}_1 + \ldots + u_n \bar{u}_n < 1 \tag{32.2}$$

ist.

Die analytische Abbildung von H' auf ein Teilgebiet von G sei durch die Gleichungen

$$x_i = c_{i1} u_1 + \ldots + c_{in} u_n + (\text{höhere Glieder}) \qquad (i = 1, \ldots, n) \tag{32.3}$$

gegeben. Da die Funktionaldeterminante

$$\Delta = |c_{ij}| \neq 0 \tag{32.4}$$

sein soll, kann man immer annehmen, daß die u_i derart numeriert worden sind, daß

$$c_{11} \neq 0 \tag{32.5}$$

ist. Durch die binäre analytische Drehung

$$\begin{cases} u_1 = \dfrac{\alpha u_1' + \beta u_2'}{\sqrt{\alpha\bar{\alpha} + \beta\bar{\beta}}}, \quad u_2 = \dfrac{-\bar{\beta} u_1' + \bar{\alpha} u_2'}{\sqrt{\alpha\bar{\alpha} + \beta\bar{\beta}}} \\ u_3 = u_3', \ \ldots, \ u_n = u_n' \end{cases} \tag{32.6}$$

wird die Hyperkugel (32.2) in sich selbst transformiert, und man muß in (32.3) die Koeffizienten c_{11} und c_{12} durch

$$c_{11}' = \frac{\alpha c_{11} - \bar{\beta} c_{12}}{\sqrt{\alpha\bar{\alpha} + \beta\bar{\beta}}}, \quad c_{12}' = \frac{\beta c_{11} + \bar{\alpha} c_{12}}{\sqrt{\alpha\bar{\alpha} + \beta\bar{\beta}}} \tag{32.7}$$

ersetzen. Nimmt man aber

$$\alpha = \bar{c}_{11}, \quad \beta = -c_{12},$$

so erhält man

$$c_{11}' = \sqrt{c_{11}\bar{c}_{11} + c_{12}\bar{c}_{12}}, \quad c_{12}' = 0. \tag{32.8}$$

Durch Wiederholung derartiger Drehungen sieht man, daß man, ohne Beschränkung der Allgemeinheit und ohne den Wert des absoluten Betrages der Funktionaldeterminante (32.4) zu ändern, die Gleichungen (32.3) auf die Gestalt bringen kann

$$\begin{cases} x_1 = c_{11} u_1 + \ldots, \\ x_2 = c_{21} u_1 + c_{22} u_2 + \ldots, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ x_n = c_{n1} u_1 + \ldots + c_{nn} u_n + \ldots. \end{cases} \tag{32.9}$$

Die Funktionaldeterminante hat jetzt den Wert

$$\Delta = c_{11} c_{22} \ldots c_{nn}. \tag{32.10}$$

Nun genügt aber, wegen unseren Annahmen, für

$$u_2 = u_3 = \ldots = u_n = 0$$

die Funktion $x_1(u_1)$ in (32.9) den Voraussetzungen des Schwarzschen Lemmas, und man folgert hieraus, daß $|c_{11}| \leqq 1$ ist. Ebenso kann man aber beweisen, daß $|c_{22}| \leqq 1, \ldots, |c_{nn}| \leqq 1$ sind, und hieraus folgt nach (32.10) die Relation

$$|\Delta| \leqq 1.$$

Durch die Gleichungen

$$x_1 = u_1, \ldots, x_n = u_n,$$

die die Funktionaldeterminante $\Delta = 1$ besitzen, geht aber die Hyperkugel (32.2) in ein Teilgebiet des Polyzylinders (32.1) über, und es gilt also der

Satz 15. *Die im Polyzylinder*

$$|x_1|<1, \ldots, |x_n|<1$$

eingeschriebene Hyperkugel

$$x_1\bar{x}_1+\ldots+x_n\bar{x}_n<1$$

ist ein Maximalteiler desselben.

33. Es sei jetzt umgekehrt G die Hyperkugel

$$x_1\bar{x}_1+\ldots+x_n\bar{x}_n<1 \tag{33.1}$$

und H' der Polyzylinder

$$|u_1|<1, \ldots, |u_n|<1. \tag{33.2}$$

Man kann, ähnlich wie in **32.**, durch analytische Drehungen, die den absoluten Betrag der Funktionaldeterminante nicht ändern, die Funktionen, durch welche der Polyzylinder (33.2) auf ein Teilgebiet von (33.1) abgebildet wird, auf die Gestalt (32.9) bringen. Die Funktionaldeterminante im Nullpunkt hat dann wieder die Form (32.10).

Nun gibt es analytische Drehungen der Hyperkugel (33.1), bei denen die erste Gleichung der Substitution die Gestalt hat

$$x_1'=\frac{e^{i\vartheta_1}}{\sqrt{n}}x_1+\frac{e^{i\vartheta_2}}{\sqrt{n}}x_2+\ldots+\frac{e^{i\vartheta_n}}{\sqrt{n}}x_n; \tag{33.3}$$

hierbei bedeuten $\vartheta_1, \ldots, \vartheta_n$ irgend welche reelle Zahlen, deren Bestimmung später erfolgen soll. Setzt man in (33.3) die Werte (32.9) ein, so erhält man

$$x_1'=c_1'u_1+\ldots+c_n'u_n+\ldots. \tag{33.4}$$

Nun seien $(\alpha_1, \ldots, \alpha_n)$ beliebige komplexe Zahlen, deren absoluter Betrag gleich Eins ist. Setzt man in die letzte Relation $u_i=\alpha_i t$ $(i=1, \ldots, n)$, so erhält man eine Funktion $\varphi(t)$, für welche alle Voraussetzungen des Schwarzschen Lemmas gelten. Es ist also

$$|c_1'\alpha_1+\ldots+c_n'\alpha_n|\leqq 1, \tag{33.5}$$

und da diese letzte Relation immer erfüllt sein muß, sobald alle $|\alpha_i|=1$ sind, hat man auch

$$|c_1'|+\ldots+|c_n'|\leqq 1. \tag{33.6}$$

Die Transformationsformeln für die c_i' lauten:

$$c_n'=\frac{e^{i\vartheta_n}}{\sqrt{n}}c_{nn},$$

$$c_{n-1}'=\frac{e^{i\vartheta_n}}{\sqrt{n}}c_{n,n-1}+\frac{e^{i\vartheta_{n-1}}}{\sqrt{n}}c_{n-1,n-1},$$

. . . . . .

Man kann daher die Zahlen $\vartheta_{n-1}, \vartheta_{n-2}, \ldots, \vartheta_1$ nacheinander so bestimmen, daß folgende Relationen gelten:

$$|c_n'| = \frac{|c_{nn}|}{\sqrt{n}}, \quad |c_{n-1}'| \geqq \frac{|c_{n-1,n-1}|}{\sqrt{n}}, \ldots, |c_1'| \geqq \frac{|c_{11}|}{\sqrt{n}}.$$

Nach (33.6) muß daher auch stets

$$|c_{11}| + |c_{22}| + \ldots + |c_{nn}| \leqq \sqrt{n} \tag{33.7}$$

sein.

Der absolute Betrag von $|\Delta|$ kann also nie größer als das Maximum von $|c_{11} \cdot c_{22} \cdot \ldots \cdot c_{nn}|$ sein, wenn die c_{ii} der Bedingung (33.7) unterworfen sind, und dieses Maximum wird erreicht, wenn man

$$c_{11} = c_{22} = \ldots = c_{nn} = \frac{1}{\sqrt{n}}$$

setzt.

Da nun andererseits durch die Gleichungen

$$x_1 = \frac{u_1}{\sqrt{n}}, \ldots, \quad x_n = \frac{u_n}{\sqrt{n}}$$

der Polyzylinder (33.2) auf ein Teilgebiet der Hyperkugel (33.1) abgebildet wird, hat man schließlich den

Satz 16. *Der in der Hyperkugel*

$$x_1 \bar{x}_1 + \ldots + x_n \bar{x}_n < 1$$

eingeschriebene Polyzylinder

$$|x_1| < 1, \ldots, |x_n| < 1$$

ist ein Maximalteiler derselben.

München, den 13. September 1931.

(Eingegangen am 15. September 1931.)